The JOURNAL *of*
EXPERIMENTAL EDUCATION

Volume IV September, 1935 Number 1

NUMBER READINESS AND PROGRESS IN ARITHMETIC

by
Gertrude Hildreth
Lincoln School of Teachers College
New York City

Failure in elementary school arithmetic and learning disability are due to many factors. Recent research activities suggest that lack of readiness for initial instruction, as it is typically organized in primary grades, may be an important contributing factor to disability. Educators and research workers are inquiring whether investigation of readiness in young learners would indicate aptitude for formal arithmetic instruction.

In order to secure a partial answer to this question, the present research project was undertaken. A threefold research program was carried on involving first, the construction and use of readiness tests in the study of pupil aptitude for primary learning; second the preparation and application of diagnostic tests suitable for measuring outcomes from primary grade instruction; and third, comparison of results from the two tests to determine the degree of relationship existing between them. Obviously the benefit to education of having reliable arithmetic prediction measures would be large. From the present study it was possible also to determine the distribution of arithmetic readiness or initial preparation for arithmetic instruction and the distribution on subsequent achievement tests for groups as a whole and for boys and girls separately. The report is part of a larger study carried on in both public and private schools to determine general readiness for primary school achievement and subsequent success in several achievement fields. This is the first of the follow-up studies that have been planned for the project.

Several previous studies have been reported of the arithmetic knowledge kindergarten and first grade children possess, among them Hall's "Contents of Children's Minds on Entering School", the investigations of Dearborn and Buckingham reported in the Twenty-ninth Yearbook of the National Society for the Study of Education, and the work of Woody, Ballard, Drummond and others. Important studies have also been made of the mental maturity required for successful arithmetic learning at various age levels. Washburne and Drummond have shown the futility of teaching children facts and skills which their immature minds are not naturally ready to learn,--facts and skills that can be more easily acquired when requisite preliminary knowledge and skill and a higher degree of mental maturation have been attained. Failure in the upper as well as the lower grades is attributable in part to essential lack of preparation in terms of number knowledge or mental maturity. In harmony with the results of these experiments, schools now tend to delay instruction until more number readiness is evidenced.

The Subjects and Tests

The subjects of the present study consisted of two groups of children in a

progressive private school enrolling children of professional parents. The children were both somewhat younger and brighter than typical first grade groups. One group entered the first grade in 1932, the other entered first grade in 1933.

The Readiness Tests

The first group was given the experimental form of the Metropolitan Readiness tests,[1] the second group the final form of the readiness tests, within the first school month.

The readiness test consists of several parts,--perception of pictures, symbols and numbers, vocabulary, words and sentences, number knowledge, and general drawing. To supplement these tests the children were asked to write their names and draw a picture of a man.

The number knowledge tests consist of size comparison, counting, ordinal numbers, writing numbers, identifying numbers, telling time, identifying fractional parts, and giving answers to simple reasoning problems such as, "If I had ten pennies and lost five how many would I have then?"

There were about one hundred children in the original groups who took the initial readiness tests in these two years, divided almost equally between boys and girls. The subjects ranged in age from five and a half years to six years, five months, with a median at approximately five years, eleven months.

A number of the original children dropped out at the end of the first and second years leaving smaller groups for the follow-up study. At the end of a year and a half for the 1933 group, two years and a half for the 1932 group, the remaining children were given achievement tests in arithmetic suited to the nature of the arithmetic instruction they had received.

Several classes of public school children were given the readiness tests at the same time, but in an attempt to check their later achievement in arithmetic it was difficult to locate even a third of the original children because of change of residence. Of those who had left their original schools the majority had scored in the lowest third or quarter on the readiness tests, and only a highly selected group remained. Consequently, data for these groups were not included in the study.

The Diagnostic Arithmetic Tests

Diagnostic achievement tests were constructed to measure typical arithmetic learning of the first, second, and third grades. The tests were partially arranged as performance material administered individually to the second grade, in part individually, and partially as group tests to the third grade. The tests contained items relating to measurement, counting, writing and reading of numbers, time, number knowledge, money recognition and computation, number combinations, fractions, and reasoning problems. The children enjoyed these tests enormously, considered them games rather than tests, and appeared to obtain real satisfaction from handling money, measuring with a yardstick, and weighing themselves. The diagnostic tests were designed to encompass the more general arithmetic and number learnings not usually measured with standard tests.

The School Arithmetic Program

The teaching of numbers in the school these children attend is largely informal in the first and second grades. In the second grade, number combinations are introduced uniformly and at the end of the year all children are expected to have mastered addition combinations under ten. Work books are introduced in the second grade but only limited use is made of them. Number situations occurring naturally in the course of classroom activities are used for instructional purposes. Children are trained in counting, telling time, estimating, adding lunch charts, and doing simple reasoning problems involving measurement or money, and reading and writing of numbers. In the third grade more formal work is done in

1. Published by the World Book Company, Yonkers-on-Hudson, New York.

mastering the subtraction and addition combinations and solving reasoning problems. Work books are used regularly and some arithmetic drill is given daily. At the same time use is made of number opportunities arising out of activity projects in the third grade for enlargement of number concepts and improvement in computation. Practice is fairly consistent from child to child, but slower learners receive additional help that may be needed.

In the first grade, too, advantage is taken of number situations for the teaching of number facts naturally, and the development of number concepts.

Comparison between Readiness Scores and Achievement Tests

The available number of cases in the follow-up study was reduced by dropping out, so that of those remaining, 26 were in the third grade, 33 in the second. The dropping out occurred equally among the more and less competent pupils. Correlations were computed between total readiness test score and arithmetic score, for these two groups separately, using the Spearman Rank formula:

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)}$$

For the third grade group the resulting correlation was .58 with a probable error of .13.

For the second grade group the correlation was .497 with a probable error of .13.

These correlations compare with other typical correlations between initial measures of abilities of various kinds and subsequent school achievement. The typical correlation is usually about .5.

Although these correlations are not high, we find on examining the rank order charts that for the extreme cases, those especially competent or incompetent in arithmetic, attainment can be predicted with a higher degree of certainty from initial readiness scores than for pupils in the middle range of the distributions. For educational guidance purposes, for promotion, classification, and grade placement, the extreme cases are the ones the school is most concerned about.

Distributions of Readiness and Arithmetic Scores

1932 Readiness Tests		1933 Readiness Tests	
Score	Number of cases	Score	Number of cases
116	1	114	1
114	3	112	
112	1	110	1
110		108	
108	1	106	1
106		104	
104		102	1
102	1	100	
100		98	3
98	2	96	
96	2	94	4
94	1	92	1
92	2	90	3
90	1	88	1
88		86	2
86		84	2
84	2	82	
82	2	80	1
80	2	78	1
78	2	76	2
76	2	74	1
74		72	1
72		70	3
70		68	2
68		66	
66		64	
64		62	
62		60	
60		58	1
58		56	
56		54	
54	1	52	
		50	
N = 26		48	
Median = 94		46	
		44	
		42	
		40	
		38	1
		N = 33	
		Median = 87.5	

Distribution of Arithmetic Test Scores

Second Grade		Third Grade	
Score	Number of cases	Score	Number of cases
94	1	94	2
92	2	92	1
90		90	1
88	1	88	4
86		86	5
84		84	2
82	1	82	2
80	1	80	
78	1	76	3
76		74	1
74	1	72	
72	1	70	1
70		68	1
68	1	66	1
66	4	64	1
64	1	62	
62	1	60	
60		58	1
58	1		
56	1	N = 26	
54		Median = 86	
52	5		
50	2		
48	1		
46	1		
44	3		
42	1		
40			
38	1		
36			
34			
32			
30			
28			
26			
24			
22			
20			
18			
16	1		
N = 33			
Median = 59			

Distribution for Boys and Girls Separately, Second Grade Groups

Readiness Test				Arithmetic Test			
No. of Boys		Girls		No. of Boys		Girls	
Score	No.	Score	No.	Score	No.	Score	No.
110	1	114	1	94	1	92	1
108		112		92	1	90	
106		110		90		88	
104		108		88	1	86	
102	1	106	1	86		84	
100		104		84		82	
98	1	102		82	1	80	
96		100		80	1	78	
94	2	98	2	78	1	76	
92		96		76		74	1
90		94	2	74		72	1
88		92	1	72		70	
86	1	90	3	70		68	
84	1	88	1	68	1	66	3
82		86	1	66	1	64	1
80	1	84	1	64		62	
78	1	82		62	1	60	
76		80		60		58	1
74		78		58		56	
72	1	76	2	56	1	54	
70	3	74	1	54		52	4
68	1	72		52	1	50	2
66		70		50		48	
64		68	1	48	1	46	1
62		66		46		44	
60		64		44	3	42	1
58		62				40	
56		60		N = 15		38	1
54		58	1	Median = 67		36	
52						34	
50	N = 18					32	
48	Median = 90.6					30	
46						28	
44						26	
42						24	
40						22	
38	1					20	
						18	
N = 15						16	1
Median = 81							
				N = 18			
				Median = 53.5			

Factors which Reduce the Correlations

A study of individual cases and of the groups as a whole indicates reasons for lack of perfect agreement between initial and final achievement scores. The following facts must be considered:

The groups were small in number.

The groups did not represent a representative cross section of the typical public school population. They are more homogeneous in initial ability. The differences in rank that appear are not so significant or so great as they would be among a broader, more general sample of the population.

Although the girls and boys were fairly equal on the initial readiness tests, the boys of the second grade group are superior to the girls on the achievement test, for no reason that is evident in the data. This situation affects the rank correlations.

Young children who made low initial readiness scores tended to gain more rapidly from instruction and to rate higher on final achievement scores. The opposite was sometimes true of duller children or those less ready, as shown by initial tests. This situation inverts rank order on the two tests for a few children.

Absence and illness accounted for change in rank in several cases.

Some of the third grade children were reaching their limit of achievement on the arithmetic tests. The inclusion of more difficult test items would have resulted in a wider distribution of scores.

Personality factors accounted for other changes. Lack of interest in learning arithmetic on the part of children originally scoring high in readiness test was observed. This was more true of girls than of boys. Girls showed more interest in reading activities.

Formal home training that anticipated the school program affected several cases.

The unreliability of the tests used affects the correlations found.

The two tests themselves were different in many respects. The readiness material contained more general tests. The arithmetic material was entirely confined to arithmetic.

The readiness tests were given to the children in groups; the arithmetic tests in whole, or in part were used as individual tests.

The initial testing of readiness was no doubt too limited. It should involve a larger investigation of children's number activities and interests, as well as paper and pencil examination of reaction to number problems and situations.

Educational Implications

This study indicates the need for more extensive arithmetic investigations of pupils in initial stages of learning. Cumulative studies are important for future learning prediction. The experiment should be repeated with larger numbers of cases in the general primary school population. More extensive study should be made of arithmetic readiness, through individual testing, and more adequate measurement of subsequent learning should follow.

Since, in the present study, the best and poorest learners could be predicted with fair reliability, the present tests can be used as rough measures of the traits in question. The extremes at least can be predicted, and the course of learning for others, if significant facts about them are considered.

Poor scorers on the readiness tests did not improve their status greatly, as the result of instruction and increased opportunity to learn, out of proportion to initial status. Equal practice did not even up achievement. For the poorest learners an even more simple and less rapidly moving instructional program is recommended. A more extensive controlled experiment with the curriculum factor varied would answer the questions of grade placement and curriculum variations for slow and bright learners.

The readiness data are important in indicating to the teacher what to expect under existing curriculum conditions, and in understanding the learning progress rate for individual children.

The arithmetic achievement tests used in the study constitute in their turn an important predictive device for future learning success in the upper elementary grades.

Comprehensive genetic studies are required to understand more fully the nature of arithmetic learning in individual children of varying abilities and interests and temperament.

RELIABILITY OF DIAGNOSIS OF CERTAIN TYPES OF ERRORS IN LONG DIVISION WITH A ONE-FIGURE DIVISOR

by
Foster E. Grossnickle
State Teachers College
Jersey City, New Jersey

It is generally known that many errors which a pupil makes in dealing with the fundamental operations are due to chance. When an error of combination is due to chance, the correct response will occur at least as often as the incorrect response. However, when a combination occurs two or more times in a test and each response to the fact is incorrect, the error is said to be constant. This study was made to determine if certain types of errors made in long division were due to chance or if they were caused by a faulty knowledge of the combinations. In the latter case the error should be constant.

A test was given in Grades V-XV, inclusive, (Grades XIII, XIV, and XV were in normal school) to determine if the long or the short form of division is better when the divisor is a one-figure number. There were approximately 200 students tested in each of the eleven grades sampled. About an equal number used the short and the long-division form. The results showed conclusively that the long-division form was superior to the short form.[1] Then, a study was made of the types of errors[2] which a pupil made when he divided in the long form with a one-figure divisor. These errors were grouped under six different classifications. The classification which had the greatest frequency of errors was known as errors of combinations. These included errors of combinations in division, multiplication, and subtraction. The errors were made when the subjects used in the experiment on short and long division solved the examples which follow.

Part I

a. $3\overline{)2844}$	f. $8\overline{)5785}$	k. $7\overline{)6595}$
b. $5\overline{)7826}$	g. $9\overline{)7589}$	l. $5\overline{)3704}$
c. $2\overline{)1714}$	h. $4\overline{)27322}$	m. $6\overline{)7838}$
d. $7\overline{)45663}$	i. $6\overline{)2761}$	n. $3\overline{)2297}$
e. $4\overline{)1546}$	j. $9\overline{)8389}$	o. $2\overline{)5900}$

Part II

A. $7\overline{)4031}$	F. $6\overline{)2031}$	K. $8\overline{)21431}$
B. $9\overline{)21211}$	G. $9\overline{)2230}$	L. $7\overline{)6124}$
C. $7\overline{)5100}$	H. $8\overline{)50300}$	M. $8\overline{)5102}$
D. $8\overline{)2215}$	I. $6\overline{)4101}$	N. $7\overline{)31352}$
E. $9\overline{)31121}$	J. $9\overline{)4220}$	O. $9\overline{)2303}$

Part I of the test was composed of examples which were easy to solve and Part II was composed of those which were difficult to solve. When reference is made to an example in Part I, a small letter will be used to designate the example; a capital letter will refer to an example in Part II of the test.

In order to determine whether an error in subtraction and in multiplication was due to chance, the responses to those subtraction and multiplication facts which were used two or more times in the test were studied. Brueckner and Ewell state that, "At least three examples of each type should be included in a test for individual diagnosis, since

1. Grossnickle, Foster E. "An Experiment With a One-Figure Divisor in Short and Long Division", Elementary School Journal, XXXIV (March, April, 1934), 496-506; 590-99.
2. Grossnickle, Foster E. "Errors and Questionable Habits of Work in Long Division with a One-Figure Divisor", Journal of Educational Research (forthcoming).

errors on simple examples are, in many cases, likely to be due to chance or mere accident."[3] Some of the facts in subtraction and in multiplication had a frequency of occurrence on the test which varied from one to five. Those facts which were used only once in these two processes were not considered in this investigation. The facts in subtraction and multiplication which were used more than once on the test are shown in Table I.

Only those facts in subtraction were included in Table I which were used two or more times in the division test and those which had a remainder greater than zero. Thus, the fact, 4-4, occurred in examples a and c, but no attempt was made to note whether errors were made on facts of this type.

The subtraction facts found in Part II of the test all bridged the decade from the units' to the tens' decade. If short division were used, the

TABLE I

FACTS IN SUBTRACTION AND IN MULTIPLICATION THAT WERE USED MORE THAN ONCE IN THE TEST OF LONG DIVISION WITH A ONE-FIGURE DIVISOR

Subtraction Facts			Multiplication Facts		
Number of Example	Subtraction Facts	Frequency	Number of Example	Multiplication Facts	Frequency
d j	3 - 1	2	e h	*4 x 3	2
a e	4 - 2	2	e h	4 x 6	2
d e g	5 - 2	3	m F F I	6 x 3	4
f k o	5 - 4	3	i I	6 x 6	2
d e	6 - 4	2	F I	6 x 8	2
b d	6 - 5	2	k L N N	7 x 4	4
h i	7 - 4	2	d k C	7 x 2	3
b l n	7 - 5	3	d A A	7 x 5	3
c f m	7 - 6	3	A C L N	7 x 7	4
f g m	8 - 6	3	C L N	7 x 8	3
a j	8 - 7	2	f D H K	8 x 2	4
j k n o	9 - 8	4	f M	8 x 3	2
C H M	10 - 4	3	H K	8 x 8	2
A O	10 - 5	2	D H K M	8 x 6	4
C H	10 - 6	2	f B G O	9 x 2	4
F H I	10 - 8	3	g j B E	9 x 3	4
B K	11 - 4	2	g E G J	9 x 4	4
A B	11 - 5	2	f D H K M	8 x 7	5
D E I K L	11 - 6	5	B E O O	9 x 5	4
B F I M N	11 - 8	5	B J	9 x 6	2
D J M N	12 - 6	4	E G	9 x 7	2
G H K	13 - 6	3			
F N O	13 - 8	3	g J	9 x 8	2
K L	14 - 8	2			

(*The first factor was the divisor in the division test.)

3. Brueckner, Leo J., and Ewell, Mary. "Reliability of Diagnosis of Error in Multiplication of Fractions", Journal of Educational Research, XXVI (November, 1932), 185.

subtraction facts identified by capital letters in Table I would not be used. In case the short form were used, higher decade subtraction would be employed. Under the latter condition, two two-figure numbers are subtracted and the remainder is a one-figure number. Thus, to find the difference between 41 and 35, the thought process is "35 from 41 are 6". However, if the long-division form is used and all of the numbers are written, the thought process for the same example would be, "5 from 11 are 6". The subtraction examples in Table I which bridge the decade were determined by the latter thought process.

ERRORS IN SUBTRACTION OCCURRING IN DIVISION

There was a possibility of making an error in twenty-four different subtraction facts which occurred two or more times on the division test. Likewise, an error could have been made in the remaining twenty facts that occurred once on the test. The errors made in the latter group of facts are not considered in this study.

The number of times a subtraction fact appeared on the test was not constant for each test paper. The fact, 3-1, was given twice on the test. However, if the pupil made an error in one of the examples in which the fact, 3-1, was found, there was a possibility that this fact would be missing from that example. On the other hand, an error in an example in which the fact, 3-1, was missing may have changed the values of the numbers in such a manner that this same fact was required to complete the solution. Thus, the total frequency of a given subtraction fact on each test was a variable. In any case in which a fact listed in Table I was found only once on a given test paper, that error was eliminated from this investigation. Likewise, any fact not given in Table I was excluded from this study. The consistency of an error in subtraction as found in long division with a one-figure divisor is shown in Table II.

Table II shows that the incorrect responses made to the subtraction facts were sporadic and not constant. In most cases, if a fact occurred three times on the division test, two of the responses were correct. There was no appreciable consistency of error in any particular grade. It was only in Grade V that the percentage of error deviated much beyond the average for the entire group. There was only a difference of 6.7 per cent between the highest and lowest percentages of error for all of the eleven grades sampled. Hence, it may be concluded that the subtraction errors were mostly accidental and due to chance at each grade level tested.

The column of totals in Table II shows that most of the incorrect responses were due to chance. However, certain pupils who used the subtraction facts incorrectly may have been consistent in their incorrect responses but in a summary table these results would have been concealed. Therefore, an analysis was made of the consistency of each response. Table III shows how often each type of response was made. From the latter table it is seen that there were only 10 cases out of a group of 426 which were consistent in their wrong responses, or slightly in excess of two per cent of the total. On the other hand, there were 390 cases out of the 426 in which only one of the possible responses was incorrect, or slightly more than ninety-one per cent of the total. Therefore, it may be stated very conclusively that a diagnosis made on the basis of the incorrect responses to the subtraction facts used in the test would have been highly unreliable, in fact very little better than a random guess.

To provide further proof that the errors given in Table II were mostly due to chance, the actual number of errors was compared with a theoretical number. The theoretical number used for comparative purposes was derived by assuming that in a purely chance error there would be only one incorrect response to a fact regardless of the frequency of the fact in the division test. Thus, if a fact occurred five times on the test, by the

TABLE II

THE ACTUAL NUMBER OF INCORRECT RESPONSES AND THE POSSIBLE NUMBER OF RESPONSES FOR DIFFERENT SUBTRACTION FACTS USED IN LONG DIVISION WITH A ONE-FIGURE DIVISOR FOR GRADES V-XV

Subtraction Fact	Grade											Total for all Grades
	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	
3-1	*-----	1-2	2-4	-----	2-4	------	------	2-4	1-2	-----	1-2	9-18
4-2	** 1-3	2-4	1-2	1-2	1-2	------	1-2	1-2	1-2	-----	-----	9-19
5-2	-----	3-9	2-6	-----	------	------	------	------	1-2	2-6	-----	8-23
5-4	-----	------	------	-----	------	------	------	------	3-10	-----	-----	3-10
6-4	1-2	------	------	-----	1-2	1-2	------	1-2	------	-----	-----	4-8
6-5	-----	------	------	-----	------	------	1-2	------	------	-----	-----	1-2
7-4	-----	1-2	1-2	1-2	2-4	------	------	------	------	-----	1-2	6-12
7-5	2-5	3-7	2-6	2-6	------	1-3	------	------	1-3	1-3	2-6	14-39
7-6	-----	1-4	1-3	-----	------	------	------	------	------	1-3	-----	3-10
8-6	1-3	1-3	3-8	1-3	6-18	1-3	2-6	------	2-6	-----	1-3	18-53
8-7	-----	------	------	1-2	------	------	------	------	------	-----	-----	1-2
9-8	-----	------	------	-----	1-4	------	------	------	------	-----	-----	1-4
10-4	2-6	2-5	2-6	4-12	4-11	1-3	3-6	------	2-6	-----	1-3	21-58
10-5	-----	------	1-2	1-2	1-2	------	------	2-4	------	2-4	-----	7-14
10-6	-----	------	------	1-2	------	------	1-2	------	------	1-2	2-2	5-8
10-8	3-4	2-6	1-3	1-3	3-9	------	1-3	------	1-3	-----	1-3	13-34
11-4	1-2	2-5	1-2	1-2	------	5-9	1-2	1-2	5-10	2-5	2-2	21-41
11-5	-----	3-5	------	------	1-3	1-2	2-2	1-3	------	2-2	-----	10-17
11-6	8-15	12-51	8-38	8-29	9-46	4-19	4-20	9-32	3-15	3-15	1-5	69-285
11-8	8-33	6-22	9-33	1-4	12-43	6-29	4-20	3-10	6-24	-----	2-10	57-228
12-6	8-22	4-17	5-14	3-11	10-35	7-26	4-16	9-28	5-20	2-8	11-38	68-235
13-6	1-3	7-15	2-6	------	4-7	4-11	3-9	4-12	6-18	3-9	1-3	35-95
13-8	10-26	11-29	6-17	1-3	17-35	7-19	4-12	3-9	6-16	3-9	2-6	70-181
14-8	2-4	1-2	1-2	------	2-4	6-10	1-2	------	1-2	4-8	1-2	19-36
Total	48-128	62-188	48-154	27-83	76-229	44-136	32-104	36-108	44-139	26-74	29-87	472-1430
Per Cent of Total	37.5	33.0	31.2	32.5	33.2	32.4	30.8	33.3	31.6	35.2	33.3	33.0

*A blank space in a column indicates that no incorrect responses were given to the fact in the corresponding row.

**1-3 means that one response out of three trials was incorrect to the fact, 4-2. In this case only one pupil made an error. For the combination 11-6, there were 8 errors in 15 trials of that combination, obviously by several pupils. Data when all response on a combination were correct are not included in the table. In similar manner each of the other numbers is to be interpreted.

TABLE III

ANALYSIS OF INCORRECT RESPONSES TO THE SUBTRACTION FACTS TAKEN FROM THE TOTALS FOR ALL GRADES AS GIVEN IN TABLE II

Subtraction Fact	Responses	Frequency	Responses	Frequency	Responses	Frequency
3-1	*1-2	**9				
4-2	1-2	8	1-3	1		
5-2	1-2	1	1-3	7		
5-4	1-3	2	1-4	1		
6-4	1-2	4				
6-5	1-2	1				
7-4	1-2	6				
7-5	1-2	3	1-3	11		
7-6	1-3	2	1-4	1		
8-6	1-2	1	1-3	17		
8-7	1-2	1				
9-8	1-4	1				
10-4	1-2	2	1-3	17	2-3	1
10-5	1-2	7				
10-4	1-2	3	2-2	1		
10-8	1-3	10	3-4	1		
11-4	1-2	14	2-2	2	1-3	3
11-5	1-2	1	2-2	3	1-3	3
11-6	1-2	3	1-3	2	1-4	5
	3-4	2	1-5	46	2-5	2
	3-5	1				
11-8	1-3	1	1-4	9	2-4	1
	1-5	31	2-5	4	3-5	2
12-6	1-2	2	1-3	7	2-3	2
	3-3	1	1-4	47	2-4	2
	1-5	1				
13-6	1-2	1	1-3	27	2-3	2
	3-4	1				
13-8	1-2	7	1-3	49	2-3	3
	3-3	1	2-4	1	3-4	1
14-8	1-2	13	2-2	2	1-3	2

*1-2 means that one response out of two was incorrect to the fact 3-1. In a similar manner each of the other numbers is to be interpreted.

**The number in the frequency column indicates how often the responses occurred in the column to the left.

hypothesis formulated, only one incorrect response should be caused by accident.

The number of chances out of 100 for spurious errors was found for each subtraction fact given in Table I. The chances for the actual number of errors for the facts given in Table II was compared with the former value. The chances of occurrence for the theoretical and actual number of errors for each subtraction fact are given in Table IV.

It is evident from Table IV that most of the errors were caused by chance. For 7 of the 24 subtraction facts, the chances for the theoretical and the actual values were the same. For 3 of the facts, the chances for the actual errors were less than for the theoretical.

TABLE IV

COMPARISON OF CHANCES OUT OF 100 OF INCORRECT RESPONSES TO SUBTRACTION FACTS IN DIVISION FOR THEORETICAL AND ACTUAL CASES

Subtraction Fact	Theoretical	Actual
3-1	50.0	50.0
4-2	50.0	47.4
5-2	33.3	33.4
5-4	33.3	30.0
6-4	50.0	50.0
6-5	50.0	50.0
7-4	50.0	50.0
7-5	33.3	35.9
7-6	33.3	30.0
8-6	33.3	34.0
8-7	50.0	50.0
9-8	25.0	25.0
10-4	33.3	36.2
10-5	50.0	50.0
10-6	50.0	32.5
10-8	33.3	38.2
11-4	50.0	51.3
11-5	50.0	58.8
11-6	20.0	24.2
11-8	20.0	25.0
12-6	25.0	29.0
13-6	33.3	37.6
13-8	33.3	38.7
14-8	50.0	52.8
Total	35.8	33.0

(This was made possible by conditions described on page 9.) Then, for 10 of the 24 facts, the chances for actual errors in subtraction facts were equal to or less than for the theoretical chances. For the remaining 14 facts, the chances of the actual number of errors was slightly greater than that of the theoretical number. From this group of 14 facts, there were 11 facts in which the chances for the actual number of errors were not increased more than 5, than the chances for the corresponding theoretical values. Thus, there were only 3 facts out of the total group of 24 in which the chances for an incorrect response were more than 5 in excess of the chances for a spurious or accidental error. These data show that practically all of the subtraction errors made in the long form of division with a one-figure divisor were sporadic. Any diagnosis made on the basis of an incorrect response to a subtraction fact occurring but once in a test would be highly unreliable. The data in Tables II-IV, inclusive, verify the statement quoted from Brueckner and Ewell that responses should be made to a fact at least three times before the results may be used to form a reliable diagnosis.

The 24 subtraction facts were divided into two equal groups of 12 facts each. Each fact in the one half ranked higher according to difficulty than the facts in the other half. The rank for difficulty was taken from the difficulty ratings of the subtraction facts by Knight and Behrens.[4] The average difference in the number of chances for error between the actual and theoretical values was found for each half of the facts. For the half composed of the easier facts, the average increase in the actual number of errors over the theoretical was 0.7; for the half composed of the more difficult facts, the average was 3.5. In each case, the average was derived from the algebraic sum of the differences. These two averages show that there was more consistency in errors among the difficult subtraction facts than among the easy ones.

ERRORS IN MULTIPLICATION OCCURRING IN DIVISION

The errors of multiplication which were found on the division test were analyzed in a manner similar to those of subtraction. The number of errors of combinations ranked second in the order of frequency of all types of errors in division.[5]

4. Knight, F. B., and Behrens, M. S. The Learning of the 100 Addition Combinations and the 100 Subtraction Combinations, pp. 29-33. New York: Longmans, Green and Co., 1928.
5. Grossnickle, Foster E. Journal of Educational Research (forthcoming), op. cit.

There were 50 different multiplication facts used on the division test but only 22 of these occurred twice or more. This latter group is given in Table I. The number of errors found in this group of 22 multiplication facts is shown in Table V.

Table V shows that most of the multiplication errors found in long division were due to chance. There was not much variation in the amount of consistency of errors among the eleven grades sampled. In Grade XI there were fewer multiplication errors than in any other grade. It was in this grade that the percentage of incorrect responses was smallest when compared with the possible number of responses. In Grade XIII the percentage between the same two values was largest. However, the difference in percentage for the two values for Grades XI and XIII was less than ten. Since there was a rather small range in variation of the ratio to correct and incorrect responses to the multiplication facts among all of the grades, it is evident that most of the errors were spurious.

The column of totals in Table V was decomposed into component parts to show how often each type of response occurred. The results of this analysis are shown in Table VI. The latter table shows that there were only 31 cases out of a group of 638 in which the incorrect responses to a multiplication fact were consistently incorrect, or slightly less than five per cent of the total. Then in about five per cent of the cases studied, a reliable diagnosis could have been made concerning the pupil's knowledge of certain multiplication facts. On the other hand, there were 506 cases out of the total of 638 in which only one response out of the various possible chances was incorrect, or slightly less than eighty per cent of the total. Therefore, at best, a diagnosis of the pupil's knowledge of the multiplication facts could be better than a random guess in only twenty per cent of the possible cases sampled in this study.

Further evidence is given to show how chance affected a pupil's response to a multiplication fact. A comparison of the number of actual errors with the number of theoretical errors was made as in the case of the subtraction facts. It must be remembered that the theoretical number of errors is found by assuming there would be one error for each fact which occurred two or more times on the division test. The number of chances out of 100 for theoretical and actual errors is shown in Table VII.

Table VII shows that for 7 of the 22 facts, the chances for the actual occurrence of errors were less than for the theoretical values. Although in none of the cases the corresponding values were the same, for 4 of the facts the chances for actual errors were not over 5 greater than for the chances of the theoretical values. For the remaining 11 facts, the number of chances for a real error was at least 10 greater than for the theoretical case. Hence, these data show that chance or accident was one of the chief factors which caused an incorrect response to a multiplication fact, when the latter was used in long division with a one-figure divisor.

The 22 multiplication facts under consideration were divided into two equal groups of 11 each. The one group was composed of the easier facts and the other group of the more difficult facts. The difficulty of a multiplication fact was determined by the ranking of the facts by Norem and Knight.[6]

As in the case of the subtraction facts, the average difference in the number of chances for error between the actual and theoretical values was found for each group of the facts. For the group of easier facts, the average chance for actual error to be in excess of the theoretical was -0.1; for the group of difficult facts, the average was 5.0. In each case, the average was derived from the algebraic sum of the differences. These two averages show that the chance was much greater for consistency in

6. Norem, Grant B., and Knight, F. B. "The Learning of the One Hundred Multiplication Combinations", The Twenty-Ninth Yearbook of the National Society for the Study of Education, pp. 555-57. Bloomington, Illinois: Public School Publishing Co., 1930.

TABLE V

THE ACTUAL NUMBER OF INCORRECT RESPONSES AND THE POSSIBLE NUMBER OF RESPONSES FOR DIFFERENT MULTIPLICATION FACTS USED IN LONG DIVISION WITH A ONE-FIGURE DIVISOR FOR GRADES V-XV

Multiplication Facts	Grades											Total for all Grades
	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	
4 x 3	*------	-------	-------	1-3	------	------	------	------	------	-----	-----	1-3
4 x 6	** 2-4	2-2	-------	2-5	------	------	------	------	------	1-2	-----	7-13
6 x 3	------	-------	-------	2-8	1-4	1-4	------	1-3	2-7	2-8	-----	9-34
6 x 6	3-8	1-3	2-6	1-2	1-3	2-4	------	------	2-6	-----	-----	12-32
6 x 8	2-4	-------	-------	1-2	2-3	------	------	------	------	-----	-----	5-9
7 x 2	1-3	2-4	3-8	------	3-6	1-3	------	------	------	-----	-----	10-24
7 x 4	4-14	5-14	2-7	------	7-10	4-11	------	------	3-7	-----	-----	25-63
7 x 5	------	4-12	7-18	2-5	1-3	------	1-3	3-8	------	-----	-----	18-49
7 x 7	8-27	5-21	6-22	5-21	4-4	6-23	2-7	5-15	3-15	-----	4-18	48-175
7 x 8	6-10	6-18	8-17	------	6-16	5-13	------	6-14	2-2	-----	1-4	40-94
8 x 2	8-25	6-21	22-61	2-9	16-45	6-22	7-23	5-17	2-9	4-13	2-8	80-253
8 x 3	2-7	2-5	6-14	4-11	2-5	3-4	------	7-16	3-7	-----	-----	29-69
8 x 6	9-22	6-25	2-6	5-10	9-31	2-8	2-9	5-16	7-24	1-4	2-8	50-163
8 x 7	12-34	15-43	15-38	5-25	5-20	6-30	2-10	4-15	6-14	5-19	-----	75-248
8 x 8	14-21	4-7	3-5	2-5	6-9	8-12	3-7	7-14	1-2	-----	4-10	52-92
9 x 2	1-4	2-9	2-9	2-9	------	------	1-5	------	------	-----	-----	8-36
9 x 3	3-10	8-28	12-33	3-14	3-13	14-32	------	3-9	4-14	-----	2-6	52-159
9 x 4	7-21	6-17	8-21	6-9	3-12	2-8	2-8	6-14	1-3	2-10	-----	43-123
9 x 5	2-9	4-14	2-8	8-19	2-9	3-6	6-23	9-24	1-4	4-14	4-14	45-144
9 x 6	8-16	10-20	10-21	3-4	11-18	11-25	6-15	9-18	7-12	4-9	6-14	85-172
9 x 7	1-2	13-19	6-16	6-14	12-25	9-20	5-13	8-20	11-18	6-13	2-4	79-164
9 x 8	7-11	6-15	4-11	2-4	3-7	3-9	3-7	5-12	6-12	-----	-----	39-88
Total	100-252	107-297	120-321	62-179	97-243	86-234	40-130	85-215	61-156	29-92	27-86	812-2205
Per Cent of Total	39.7	36.0	37.4	34.6	39.6	36.7	30.8	38.6	39.9	31.5	31.4	36.8

*A blank space in a column indicates that no incorrect responses were given to the fact in the corresponding row.

**2-4 means that two responses out of four were incorrect to the fact, 4 x 6. Data for pupils who made no errors on this combination are not included. In a similar manner each of the other numbers is to be interpreted.

TABLE VI

ANALYSIS OF INCORRECT RESPONSES TO THE MULTIPLICATION FACTS TAKEN FROM THE TOTALS FOR ALL GRADES AS GIVEN IN TABLE

Multiplication Fact	Responses	Frequency	Responses	Frequency	Responses	Frequency
4 x 3	*1-3	**1				
4 x 6	1-2	4	2-2	1	1-3	1
6 x 3	1-3	3	1-4	5	1-5	1
6 x 6	1-2	4	1-3	8		
6 x 8	1-2	3	2-3	1		
7 x 2	1-2	1	1-3	2	1-4	1
	2-4	3				
7 x 4	1-3	5	2-3	1	1-4	7
	2-4	2	4-4	1	3-5	1
7 x 5	1-2	2	1-3	11	1-4	1
	2-4	2				
7 x 7	2-2	1	1-3	6	1-4	14
	2-4	3	4-4	1	1-5	15
	1-6	1				
7 x 8	1-2	5	2-2	3	1-3	11
	2-3	3	1-4	7	2-4	1
	3-4	1				
8 x 2	2-2	1	1-3	5	1-4	26
	2-4	3	3-4	1	1-5	11
	2-5	9	3-5	2	3-6	1
8 x 3	1-2	5	2-2	1	1-3	11
	2-3	1	1-4	1	2-4	2
	3-4	1	1-5	1		
8 x 6	1-3	4	2-3	1	1-4	26
	2-4	3	1-5	3	2-5	1
	3-6	1	4-6	1		
8 x 7	1-3	3	2-3	1	1-4	3
	3-4	2	4-4	1	1-5	30
	2-5	3	3-5	1	2-6	4
	3-6	2	4-6	1		
8 x 8	1-2	21	2-2	4	1-3	8
	2-3	3	3-3	3		
9 x 2	1-4	4	1-5	4		
9 x 3	1-3	2	2-3	1	1-4	12
	2-4	6	3-4	2	1-5	10
	2-5	4				
9 x 4	1-3	1	2-3	4	1-4	15
	2-4	1	3-4	1	1-5	6
	4-5	2				
9 x 5	1-3	4	2-3	2	3-3	2
	1-4	19	2-4	1	1-5	6
	2-5	2				
9 x 6	1-2	34	2-2	2	1-3	23
	2-3	6	3-3	3	3-4	1
9 x 7	1-2	35	2-2	5	1-3	23
	2-3	1	3-4	3		
9 x 8	1-2	15	2-2	2	1-3	16
	2-3	2				

*1-3 means that one response out of three was incorrect to the fact, 4 x 3. In a similar manner each of the other numbers is to be interpreted.

**The number in the frequency column indicates how often the responses occurred in the column to the left.

TABLE VII

COMPARISON OF CHANCES OUT OF 100 OF INCORRECT RESPONSES TO MULTIPLICATION FACTS IN DIVISION FOR THEORETICAL AND ACTUAL CASES

Multiplication Fact	Theoretical	Actual
4 x 3	50.0	33.3
4 x 6	50.0	53.8
6 x 3	25.0	26.4
6 x 6	50.0	37.5
6 x 8	50.0	55.6
7 x 2	33.3	41.6
7 x 4	25.0	39.7
7 x 5	33.3	36.8
7 x 7	25.0	27.8
7 x 8	33.3	42.5
8 x 2	25.0	31.6
8 x 3	50.0	42.0
8 x 6	25.0	30.7
8 x 7	20.0	30.2
8 x 8	50.0	56.5
9 x 2	25.0	22.2
9 x 3	25.0	32.7
9 x 4	25.0	35.0
9 x 5	25.0	31.2
9 x 6	50.0	49.5
9 x 7	50.0	48.2
9 x 8	50.0	44.3
Total	32.4	36.8

error among difficult facts in multiplication than among the easier ones. This situation may indicate that the more difficult facts, not being well learned tend to have associated with them certain incorrect responses. These findings therefore reinforce the principle that initial instruction is the important thing in teaching number.

SUMMARY

This study was made to determine the consistency of an incorrect response to a basic fact in subtraction and in multiplication. The facts were used in long division with a one-figure divisor. An analysis was made of incorrect responses to 24 subtraction facts and 22 multiplication facts. The subjects used in this investigation were pupils in Grades V-XV, inclusive. The following results were obtained:

1. In about 91 per cent of the cases in subtraction, only one of the possible responses to a fact was incorrect.

2. In about 80 per cent of the cases in multiplication, only one of the possible responses to a fact was incorrect.

3. There was a greater tendency for an error to be consistent for the difficult facts than for the easier facts.

4. A perfectly reliable diagnosis concerning a pupil's knowledge of the basic facts could be made in only about two per cent of the cases for the subtraction facts and about five per cent of the cases for the multiplication facts.

5. For diagnostic purposes a pupil must be given the opportunity to make at least three responses to each basic fact in subtraction and in multiplication when this fact is used in long division, when the divisor is a one-figure number.

- - - - -

AN EXPERIMENTAL STUDY OF TWO PLANS OF SUPERVISED STUDY IN FIRST YEAR ALGEBRA

by
Burton J. Stallard
Junior High School
Wauwatosa, Wisconsin
and
Harl R. Douglass
University of Minnesota

The investigation reported here was planned and carried on in an effort to throw some light on the question of whether the long unit assignment plan of supervised study is superior to the divided period daily assignment plan. Two pairs of equated sections of ninth grade algebra at the Wauwatosa, Wisconsin, Junior High School were employed. There were 24 subjects in each section. One pair was made of superior boys and girls with I.Q.'s of from 105 to 131, mean 115.6, and the other pair of average pupils with I.Q.'s from 93 to 118, mean 105.6.

The experimental variables.- One of each of these pairs of sections was taught by use of large unit assignments covering on the average about two weeks. Mimeographed assignments and instruction sheets were distributed to each pupil. Two full class periods ordinarily were spent in discussing the material in the assignment. Instructions were given, questions were asked, explanations made, and demonstration work done on the blackboard. On the third day the pupils started to prepare the work of the assignment, and spent the entire period in supervised study. The teacher moved about the room from one pupil to another throughout the period giving assistance where it was deemed necessary. At the end of the period the assignment sheets, together with all written work, were collected, to be re-distributed the following day. This procedure was repeated each day until the stipulated time for the assignment had elapsed. At the end of each assignment period, a test was given covering the work of the unit, and one period was devoted to general remedial work before the new assignment sheets were passed out.

In each of the other sections, each daily period was divided: the first half of the period was devoted to recitation, discussion and assignment activities, and the last half of the period to supervised study. The division of the period was carefully watched and strictly adhered to throughout the experiment. No subdivisions were strictly followed in the recitation half of the period, but approximately two-thirds or twenty minutes of the period were given to oral discussion, question and answer activity, and demonstration work leading up to the next assignment. The last ten minutes were quite definitely limited to the assignment and to suggestions for procedure.

During the last thirty minutes of the period the pupils prepared the assignment given. The teacher spent the full time in passing from one pupil to another, taking note of pupil progress and giving individual suggestion wherever it appeared necessary. Good order was maintained at all times, and wherever individual assistance was given it was done in a quiet tone of voice, not disturbing to the other pupils. All papers were collected at the end of the period, whether the assignment had been completed or not. In most instances the assignment was completed. The teacher

took particular care to limit the assignment so as to be within the reach of all, and added several additional problems in each case under the title of "extra credit" work, which kept each individual occupied the full time.

Holding nonexperimental variables constant.- The pupils who were the subjects of the experiment in each pair of sections were so selected as to be of approximately equal mean ability and of approximately equal variability. In other words, the two sections of each pair were so made up as to be of approximately equal means and standard deviations in chronological age, intelligence quotients (Otis Self-Administering Tests of Mental Ability), and Standard Achievement Tests in Arithmetic (Form W as Indicated in Table I).

No pupil was included in the experiment who was absent more than nine days. There was no significant difference in the average absence of the paired sections. There was in effect a plan of make-up work whereby students after school made up work missed during absence, and absent pupils were not permitted to fall behind in school work.

The superior pair of sections met at 12:30 and 1:30, and the other pair at 8:30 and 10:30 A.M. If one section lost time due to fire drills, announcements, or early dismissals, the papers were collected in the other section a few minutes earlier to correspond.

In every way, efforts were made to hold everything constant except the experimental variables, namely, the plan of supervised study.

TABLE I

DATA CONCERNING COMPARABILITY OF PAIRED SECTIONS

Superior Pair	Mean		Standard Deviation	
	Daily Assignment	Large Unit	Daily Assignment	Large Unit
Age	14.4	14.5	6.2	6.6
I.Q.	115.6	115.8	6.5	6.1
Arithmetic	104.0	103.8	5.7	5.8
Average Pair				
Age	14.8	14.7	6.8	5.6
I.Q.	105.6	105.6	5.9	5.5
Arithmetic	97.0	96.8	8.7	8.2

In pairing the sections boys were paired with boys and girls with girls, hence the two sections had the same proportion of each sex. No homework was assigned in either class. The same teacher taught both sections. The same text was used by both sections.

Measurement of results.- As a means of measuring progress, the Douglass Standard Survey Test was given at the beginning and at the end of the experimental period of a semester in length. The gains in scores on this test were taken as a measure of progress. The coefficient

of reliability of this test was estimated by the chance-half Spearman-Brown technique to be .756.

Of the two superior sections, the long unit group gained a total of 197 points as compared with 170 by the daily divided period group, an average superiority per pupil of 1.1 points or 16%, which was 1.90 times the standard error of the difference (.60).

Of the two average sections, the daily divided period group had a slightly superior average gain per pupil of .42 points or 7%, only .95 times as great as the standard error of the difference and hence not statistically significant.

When compared on the basis of the per cent gained of the amount of possible gain, the results were similar to those obtained by other methods. The long unit section was the superior of the two superior groups by 21.5 per cent, there being no significant difference between the two average sections, with a slight but unreliable difference in favor of the daily recitation section.

These sections were tested again at the end of the second semester, during which they had been regrouped in other sections and taught by various teachers with various methods. There was no difference between the average scores of the two superior sections on the final second semester tests. Of the two average groups, the daily recitation method was superior by an insignificant amount, only .21 of the standard of the difference.

Conclusions.- The type of large unit plan of supervised study employed by Stallard was found to be superior to his plan of daily recitation plan in sections composed of superior pupils, though no significant difference was found between the two plans for sections of pupils of average ability. Until more complete evidence is available, presumption should be in favor of the long unit plan for superior pupils only.

These results are at slight variance with the findings of other investigators, who, with the exception of Hunziker, find results favorable to large units even with average or representative sections. Williams found with sections in second year algebra that a modified form of the Winnetka contract plan was superior. Hunziker found negligible differences in favor of larger unit assignments in algebra and definite differences in favor of daily units in geometry. Gadske found unmistakable favorable differences in favor of longer units. Both of the studies were carried on with representative sections of first year algebra. Eilberg and Hare in geometry found the large unit method superior with representative groups. It is quite reasonable to give the large unit method the benefit of the doubt, since five instructors have attained the same relative degree of skill in its use that they have in the use of older methods.

BIBLIOGRAPHY

Eilberg, Arthur. "The Dalton Plan versus the Recitation Method in the Teaching of Plane Geometry." Doctor's Thesis, Temple University, Philadelphia, 1931.

Gadske, R. E. "A Comparison of Two Methods of Teaching First Year High School Algebra", School Science and Mathematics, XXXIII (June, 1933), 635-40.

Hare, John S. "An Experimental Study of Two Types of Teaching Procedure in Geometry." Master's Thesis, Ohio State University, 1923.

Hunziker, Charles W. "A Comparison of the Relative Effectiveness of a Plan of Supervised Study and the Daily Recitation Method in the Teaching of Algebra and Geometry." Master's Thesis, University of Minnesota, 1933.

Williams, G. B. "A Controlled Experiment to Determine the Efficiency of the Contract Method of Teaching Second Year Algebra to Normal and Superior Pupils." Master's Thesis, Penn State College, 1932.

AN EXPERIMENTAL EVALUATION OF A MODIFIED MORRISON PROCEDURE IN TEACHING AMERICAN HISTORY

by
Harl R. Douglass
University of Minnesota
and
Kenneth L. Pederson
Hibbing, Minnesota, High School

The Problem.- Supervised study, first formally recognized as a distinct classroom procedure in the form of a double period, the first half given to recitation and the second to study of the daily lesson assigned for the morrow, for administrative reasons passed quickly into the stage in which a period of from 55 to 65 minutes was divided daily into two phases, recitation and supervised or directed study. More often than not recitation preceded study, but particularly in recent years the S-R (study-recitation) sequence has been employed frequently. Practice has also been divided between a set daily division of the class period into two sections of standard length, for example, 30 minutes recitation and 25 study and a flexible arrangement whereby the proportions were at the direction of the teacher. Whatever form was followed, if it involved the R-S sequence, teachers tended unfortunately to employ more time for recitation and less for supervised study than was contemplated, and supervised study fell into bad repute in many localities. Where the S-R sequence was employed, still other obstacles, which need not be mentioned here, were encountered.

In the last decade, plans of supervised study involving assignments or "units" of more than one day have come to be used by an increasing number of teachers. Along with the project method, Dalton, Winnetka, and Wisconsin plans in various forms and modifications, the Morrison organization has become somewhat popular.

Experimentation for the purpose of determining the relative effectiveness of large unit plans of supervised study and daily unit plans, whether or not supervised study was used, has not produced consistent results. From the experimental studies listed in the bibliography at the close of this report, the following conclusions seem justified. There is, in general, probably little difference in the merit of the two types of plans, but results favoring large unit plans seem to have been found more frequently than the contrary. With classes in history and civics, Esson and Cole, Willard, Wrinkle, Phipps, Mason, Stegmeir, Steiner, Weeks, and Slagle found the large unit method superior to the daily supervised study or daily recitation plan, though the differences in the experiments considered separately were usually quite small and statistically insignificant. Deam, Shepard, and Funk found no differences worth mentioning, and McClure found results favoring the daily plan.

It was the purpose of this investigation to furnish additional light on the general problem of large versus daily units, and on the specific problem of the relative effectiveness of a modified Morrison procedure as compared with the divided period S-R organization of supervised study.

The Subject of the Experiment.- For experimental purposes, eight sections of American History in the Hibbing, Minnesota, High School, a three year senior high school enrolling approximately 1100

pupils were used. Three of these were taught by one teacher and two units of the experiments involved these sections, Section B being paired and compared with Section A and Section C respectively. All of the remaining five sections were taught by another teacher, Sections D and E, E and F, and G and H forming three other units of the investigation.

The experimental procedure was applied during the first semester for a period of 54 actual teaching days and again during the second semester for a period of 52 days. In the second semester an arrangement was employed identical with that described in the preceding paragraph. Ninety pairs of pupils were studied in the first semester and eighty-eight in the second semester. While pupils were rematched in pairing for the second semester, no pupils were included who had changed from an experimental to a control section or vice versa.

The Experimental Variable.- In those classes taught by the daily S-R plan, the first thirty minutes were given to study and the last twenty-seven minutes to supervised study. This daily division was observed rather closely.

In the so-called "Modified Morrison" plan, the course of study was divided into units of from three to six days each. Each unit was begun with a preview of the materials to be covered and the introduction of the assignments, including suggestions on how to attack them and efforts to arouse interest in the unit. This ordinarily required a full class period or somewhat less and was followed by several days of study, concluded by written tests, oral discussions, and reports or both. Assimilation tests were not employed. The pupils of all sections were supplied with the same mimeographed "essential" sheets, outlines, study questions, and suggestions for extra credit work. The essential sheets contained lists of important dates and personages, words to define, problems, and questions. The extra-credit sheets gave the requirements for marks of C, B and A, and included special work reports and projects by which the pupils could work for the higher grades. In all classes a list of suggestions on how to study were supplied to the pupils and the contents thereof emphasized.

Control of Nonexperimental Variables.- In all five units of the experiment, the same teacher taught both paired sections. Both teachers were experienced in both methods of study and apparently were not prejudiced in favor of either method. Paired sections met during adjacent periods of the day. Pupils absent more than six of the fifty-four days involved in the experiment were not included in the investigation.

In order to reduce the possibility of serious influence of systematic error, a careful attempt was made to include data only from such pupils in each pair of sections as would result in evenly matched experimental and control groups. Pupils were paired individually on the basis of I.Q. (Otis Self-Administering Test of Mental Ability), age, final mark in a previous history course, and an initial achievement test in American History covering the work of the semester. While boys were not paired with boys or girls with girls in more than three fourths of the cases, there were approximately the same proportions of pupils of the two sexes. Data relative to the pairing are shown in Table I.

Measurement of Learning.- Three measures of relative progress were employed: (1) The Iowa Every Pupil Test in American History, an objective test prepared under the auspices of the School of Education of the University of Iowa and employed in the state "brain" contest sponsored by that institution, and (2) two objective achievement tests based on just the materials studied during the two semesters respectively. The latter tests were constructed by Mr. Pederson, one of the authors of this report, and the appropriate one was given at the beginning and at the end of each experimental period.

The coefficients of reliability

TABLE I

I.Q., AGE, PREVIOUS HISTORY MARK, AND INITIAL TEST SCORE MEANS OF EXPERIMENTAL AND CONTROL SECTIONS

	I.Q.		Age in Months		Previous History Mark		Initial American History Test	
	Control	Exper.	Control	Exper.	Control	Exper.	Control	Exper.
First Semester:								
Unit I, N = 18	103.9	104.4	197.0	198.4	81.0	81.4	26.5	27.4
II, N = 20	105.4	105.5	194.7	194.4	81.7	82.3	27.6	26.4
III, N = 20	105.3	105.3	194.0	195.4	82.8	82.9	32.5	31.1
IV, N = 20	104.4	103.8	198.4	199.6	82.2	81.6	33.4	33.5
V, N = 12	106.6	106.9	196.0	194.4	84.4	84.2	33.7	32.7
All, N = 90	105.0	105.1	196.0	196.7	82.3	82.4	30.6	30.1
Second Semester:								
Unit I, N = 18	103.6	103.7	197.8	198.9	81.3	80.8	38.7	40.0
II, N = 20	104.3	104.5	195.3	194.7	81.2	82.0	39.6	39.0
III, N = 18	105.5	105.3	193.5	194.8	82.8	83.1	51.4	51.8
IV, N = 20	103.7	103.7	198.1	198.1	81.9	81.4	49.6	50.3
V, N = 20	107.3	108.0	196.4	193.2	83.2	85.9	48.2	47.2
All, N = 88	104.6	104.8	196.2	196.1	82.1	82.3	45.3	45.6

TABLE II

AVERAGE GAINS OF SECTION M-M AND OF S-R SECTIONS ON PEDERSON TEST

	M-M	S-R	Diff.	% Diff.	Reliability of Diff.
First Semester:					
Unit I	60.9	59.0	1.9	3.3	73 in 100
II	66.2	60.1	6.1	10.2	98 in 100
III	48.4	44.2	4.2	9.5	93 in 100
IV	42.6	43.5	-.9	-2.1	62 in 100
V	49.3	47.7	1.6	3.3	74 in 100
All	53.7	51.0	2.7	5.3	95 in 100
Upper half	58.2	53.5	4.7	8.8	98 in 100
Lower half	49.2	48.5	.7	1.4	61 in 100
Second Semester:					
Unit I	62.3	59.6	2.7	4.6	74 in 100
II	63.7	62.7	1.0	1.7	61 in 100
III	48.1	43.4	4.7	10.7	99 in 100
IV	47.3	49.4	-2.1	4.4	75 in 100
V	54.5	51.6	2.9	5.7	76 in 100
All	55.3	53.6	1.7	3.1	84 in 100
Upper half	59.2	59.8	-.6	1.0	60 in 100
Lower half	51.3	47.4	3.9	8.5	96 in 100

of these tests, calculated by the odds vs. evens and Spearman-Brown techniques, were as follows: The Iowa Every Pupil Test, first semester, .83; the Pederson, first semester test, .94; and the Pederson second semester test, .90. For additional evidence, three "unit" tests, prepared by Mr. Pederson, were employed during the first semester and two during the second semester.

As measures of pupil learning, the gains on the Pederson tests were taken as the principal criterion, and the final test scores on the Iowa and the unit tests were employed as accessory data.

Relative Effectiveness of the Two Plans.- As shown in Table II, the sections employing the Modified Morrison Plan the first semester, for four of the five units, as a whole, made slightly greater gains on the Pederson test than did their pairs in S-R daily sections. The differences in gain in each instance were relatively small, ranging from 2 to 10% of the gain of the less favored section, and no one difference was statistically significant--that is, absolutely unattributable to chance errors of sampling and measurement, though, as a matter of fact, the difference in Unit II would be found merely by chance only 2 times in 100, that in Unit III only 7 in 100, and that in the total of all five groups only 5 in 100.

On the three first semester "units" tests, the differences favored the two methods impartially, eight of the fifteen favoring the S-R procedure and the other seven favoring the M-M procedure. The larger differences were, however, those in favor of the S-R plan.

On the Iowa Every Pupil Test in three of the units differences favored the M-M method, and in two the S-R sequence; while on the whole, a very slight difference in favor of the M-M procedure was found. Marks assigned by instructors favored M-M sections slightly (.25 to 2.40 points on a scale of 100) in three of the five groups and very slightly (.4 of a point) for the 90 pairs of pupils as a whole.

Second Semester.- The results for the second semester were quite similar (Table II). In four of the five units, the gains on the Pederson test favored the M-M plan as did the gain of the entire 88 pairs taken as a group. These differences ranged from less than 2 per cent to approximately 10% of the gains of the inferior section. The only difference approaching statistical significance was that found in Unit III favoring the M-M plan, which could be attributable to chance errors of measurement and sampling only once in 100 times, although the difference for the whole 88 pairs (also favoring the M-M method) might be found by chance only 16 times in 100.

For all ten comparisons on the two "unit" tests, the differences favored the M-M section, and in four of the five groups the differences on the Iowa tests favored the M-M sections. Similarly the marks given by instructors were on the average slightly higher for the M-M pupils in four of the five pairs of sections and for the 88 pairs of pupils as a whole.

Relative Merit of the Two Procedures for Bright and for Slow Pupils.- At the conclusion of the first semester, the gains on the Pederson test and the final scores on the Iowa and unit tests were studied for the upper half of the pupils and for the lower half, differentiated on the basis of I.Q. The results were quite clean cut, indicating the distinct superiority of the M-M method for the upper half of the pupils and the relatively similar effectiveness of the two methods for pupils of lower than average intelligence. On the Pederson test the average gains of the 45 bright M-M pupils were greater than those of the 45 bright S-R pupils by an amount equal to about 9% of the average gain of the latter.

The data of the second semester are not so consistent--greater gains being made on the Pederson test by the 45 slow M-M pupils as compared with the 45 slow S-R pupils, while no material difference could be noted between the two

methods for the bright pupils. On the other hand, the results of the Iowa final test indicate a distinct superiority of the M-M plan for the bright pupils and no material difference between the plans for the slow pupils.

Teachers' Opinions.- Sixteen teachers of history and social studies in the Hibbing High School, all experienced and familiar with the M-M procedure, were asked for their opinion of the relative effectiveness of the two plans. None believed that the M-M plan would be more effective for the entire group of pupils, although two thought that there would be no difference and eight thought it might be better for bright pupils.

Summary of Conclusions.- From the data collected in these investigations, it seems reasonable to conclude that:

1. The unit plan described earlier in the article, in the hands of well-trained teachers familiar with the plan, is likely to contribute to results slightly superior to those obtained by a study-recite plan of supervised study--at least as judged by the results of objective tests and teachers' marks in classes in which objective tests are employed.
2. The unit plan in question is probably better suited for bright than for dull pupils.
3. Teachers may not be relied upon to predict accurately the relative effectiveness of two teaching procedures so nearly equivalent as are the two plans of supervised study employed in this investigation.

BIBLIOGRAPHY

Douglass, H. R. Modern Methods in High School Teaching. Boston: Houghton-Mifflin Company, 1926.

Morrison, H. C. The Practice of Teaching in the Secondary School. Chicago: University of Chicago Press, 1931.

BULLETINS AND MONOGRAPHS

Douglass, H. R. "The Experimental Comparison of the Relative Effectiveness of Two Sequences in Supervised Study," University of Oregon Publications, Education Series I, No. 4, Eugene, Oregon: University of Oregon Press, 1927.

Monroe, W. S., and Others. "Special Methods on the High School Level," Review of Educational Research, I (February, 1932).

Monroe, W. S., and Others. "Psychology and Method in the High School and College," Review of Educational Research, IV (December, 1935).

Phipps, M. "The Unit Mastery Method Versus the Daily Recitation Method in Teaching High School History," Department of Secondary School Principals Bulletin, XXIV (January, 1929), 105-07.

Steiner, H. R. "A Comparative Study of the Long and Short Assignment," National Education Association, Department of Social Studies, Second Yearbook, 1932.

PERIODICALS

Deam, T. M. "Unit Organization Versus Chronological Organization in Teaching American History," School Review, XXXVIII (December, 1930), 782-86.

Esson, V., and Cole, R. "The Effectiveness of the Contract Method as Compared with the Ordinary Method of Teaching," School Review, XXXVII (April, 1929), 272-81.

Freeble, C. R. "An Experiment in Directing Thinking in Modern European History," University of Pittsburgh School of Education Journal, V (March, 1930), 97-100.

Funk, M. N. "A Comparative Study of the Results Obtained by the Method of Mastery Technique and the Method of Daily Recitation Assignment," School Review, XXXVIII (May, 1928), 388-45.

Kimmel, W. K. "Trends in the Teaching of History," Historical Outlook, IV (April, 1929), 180-84.

Morrison, H. C. "The Unit Mastery Technique as Applied to the Teaching of History," Historical Outlook, XXIV (January, 1935), 19-21.

Shepard, E. L. "Contract Versus Traditional Method in Teaching Sixth Grade History," University of Pittsburgh School of Education Journal, V (December, 1927), 47-51.

Thompson, W. H. "An Experiment with the Dalton Plan," Journal of Educational Research, XXVI (March, 1933), 493-500.

Willard, M. "An Experiment in the Use of Two Methods of Instruction," Journal of Educational Method, VIII (June, 1929), 505-10.

Wrinkle, W. L. "The Relative Merit of the Whole and the Part Method in the Teaching of the Social Studies," The Historical Outlook, XXII (November, 1931), 338-41.

UNPUBLISHED THESES

Mason, H. C. "A Comparison of the Contract Assignment and the Daily Assignment Methods of Instruction as Measured by Objective Tests," Unpublished Master's Thesis, University of Minnesota, 1930.

McClure, J. P. "An Experimental Study of the Effects of Two Plans of Distributing Study and Discussion in the Unit Method of Teaching History," Abstract of Unpublished Master's Thesis, University of Colorado, 1930.

Slagle, L. M. "Relative Value of Recitation and Laboratory Activities in Social Sciences," Abstract of Unpublished Master's Thesis, University of Southern California, 1926.

Stegmeir, C. C. "An Experimental Study of the Daily Recitation and the Mastery Technique of Teaching," Abstract of Unpublished Master's Thesis, University of Chicago, 1930.

Weeks, J. K. "Effectiveness of Two Methods of Teaching History," Abstract of Unpublished Master's Thesis, George Peabody College for Teachers, 1931.

- - - - -

THE TRANSFER OF TRAINING IN HIGH SCHOOL LATIN TO ENGLISH GRAMMAR, SPELLING, AND VOCABULARY

by
Harl R. Douglass
University of Minnesota
and
Clifford Kittelson
Ray, North Dakota, High School

Proponents of Latin as a subject in the curriculum of secondary schools have included among the valuable outcomes of teaching Latin the fact that by study of the language, students would experience growth in the mastery of English grammar, spelling, and vocabulary. The doubters have either denied these claims or insisted that the increments of growth attributable to the study of Latin are indeed small, if not negligible, and disproportionately small in comparison with time spent on the study of Latin in order to achieve them.

Since growth in the mastery of these acquisitions does not submit readily to casual observation or even to the most accurate means of measurement available, until the last two decades, the issue has only within recent years been attacked by methods other than speculation, most frequently partisan, and therefore inviting little confidence.

Previous Investigations

As early as 1915 Harris[1] reports a study concerning itself with the effect of Latin on English spelling and rhetoric. He gave tests to college freshmen in rhetoric classes, and the results in spelling showed that pupils with no Latin, one, two, and three years of Latin scored about the same, while pupils who had four years of Latin were distinctly superior. He had no measure of the initial ability of these people, however, and later studies would tend to indicate that the four-year Latin pupils were of superior initial ability.

Foster[2] likewise gave a test in spelling, composed of Latin derived words from Lodge's Spelling List. He reports some superiority for three- and four-year Latin pupils, but upon checking intelligence he discovered both that Latin pupils were mentally superior, and that there was a close relationship between mental ability and spelling ability. He concluded that Latin, as it was being taught, was of little value in spelling, and that spelling ability was gained by a pupil previously to his taking Latin.

Dallam[3] equated groups on the basis of average scholarship, and gave tests in spelling, dictation, Latin derivation, definition, composition, and English grammar. The only tests in which she found Latin pupils distinctly superior were in derivation and grammar.

These earlier studies suggested that Latin was not contributing as much as had often been claimed, and that the fault might lie partly in content and teaching methods. The feeling of a need for more thorough investigation of this entire problem resulted in a concerted effort on the part of the American Classical League to promote a

1. Harris, James H., "A Study in the Relation of Latin to English Composition," School and Society, II (August, 1915), 251-52.
2. Foster, F. M., "The Results of a Recent Spelling Test at the University of Iowa," School and Society, V (April, 1917), 506-08.
3. Dallam, M. F., "Is the Study of Latin Advantageous to the Study of English?" Educational Review, LIV (December, 1917), 500-03.

thorough investigation on a nation-wide scale. The resulting Classical Investigation had two main objectives, namely, to determine objectively the value of Latin as it was being taught in the average high school, and, through experimental procedures, to make recommendations as to how Latin could be taught more effectively to achieve the claims made for it.

Contributing to the Classical Inquiry, perhaps the outstanding derivative studies are those of Hamblen[4] and Haskell,[5] both working for doctorates at the University of Pennsylvania. Hamblen found that method played an important part in the results achieved. He shows that the amount of automatic transfer from Latin to knowledge of English vocabulary is very slight. He goes on to show that if derivation is taught in Latin classes, without developing rules and principles as aids, transfer increases. In classes where derivation was studied with the aid of rules and principles still better results were achieved.

Haskell's conclusions support those of Hamblen. In studying the effect of teaching derivation in English classes, he reports that the results are about equivalent to automatic transfer in Latin classes where no attention is paid to derivation.

These investigators carried on their work in Philadelphia high schools, basing their results on the data gathered in ninth grade English and Latin classes, in which the experimental groups spent approximately one-fifth of their time in the study of derivatives. They paired the pupils on the basis of intelligence tests in the form of the Terman Group Test of Mental Ability. Haskell also used the initial results of two standardized vocabulary tests in pairing.

Thorndike and Ruger,[6] in an investigation involving more than 5,000 pupils, studied the effect of Latin on English vocabulary. They base their conclusions on the results of the Carr Test of English vocabulary, administered in three forms during the course of a year. Their pairings are on the initial scores of the Carr Test. Their figures indicate that during the course of a year Latin pupils gained about two and a half times as much in vocabulary of Latin origin as did non-Latin pupils. In words of non-Latin origin the two groups gained about equally.

The last three studies have considered only ninth grade pupils, and the results produced during the course of one school year. Otis[7] paired two-year Latin pupils with pupils who had not studied Latin, on the basis of intelligence, and gave tests in vocabulary and composition. He found that Latin pupils excelled non-Latin pupils by thirty-three per cent in vocabulary, and six and a half per cent in composition.

A valid criticism, applicable to most of these studies, is made by Orleans,[8] when he demonstrates that a great number of words of Latin derivation have undergone marked changes, often to the extent that the reader is not aided in reasoning out the meaning of a word with the background which high school Latin gives. The Latin derived words in the tests of these studies are mostly taken from words common to the derivative lists of first-year Latin textbooks, or at least common to derivative lists

4. The Extent to Which the Effect of the Study of Latin upon a Knowledge of English Derivatives Can be Increased by Conscious Adaptation of Content and Method to the Attainment of this Objective. Philadelphia: University of Pennsylvania Press, 1925. Pp. 81.
5. The Comparative Results Produced by Teaching Derivation in the Ninth Grade Latin Classes and in the Ninth Grade English Classes of Non-Latin Pupils in Four Philadelphia High Schools. Philadelphia: University of Pennsylvania Press, 1923. Pp. 139.
6. Thorndike, Edward L. and Ruger, G. J., "The Effect of First-year Latin upon a Knowledge of English Words of Latin Derivation," School and Society, XVIII (September, 1923), 260-70.
7. Otis, Alvah T., "The Relation of Latin to the Study of English Vocabulary and Composition," School Review, XXX (January, 1922), 45-50.
8. Orleans, Jacob S., "The Possible Transfer of a Study of Latin to English Vocabulary," School and Society, XVI (November, 1922), 559.

in the typical four-year high school course. The results, therefore, have been presented in terms of the percentage of Latin derived words in English, while the tests upon which the results are based were not a representative sample of Latin derived words.

Coxe[9] studied the gains seventh and ninth grade pupils made during the course of a year in spelling, when studied under different conditions in both Latin and English classes. He reports that, in Latin classes where no provision was made for study of spelling, this group gained about as much in Latin derived words as did the English classes where rules were developed to aid in spelling. The Latin group, in which about one-fifth of the time was spent in studying spelling without developing rules, gained more than the above two groups. Superior to all these groups were the classes in which principles and rules had been developed to aid in spelling.

Miller and Briggs[10] studied the effect of Latin translations on English, gathering their data from written and oral translations. They report that thirty-four per cent of the translations did not convey any meaning, that forty per cent of the translations were awkward, while only six per cent were in "impeccable English." Analysis of these translations showed that a great many errors were due to disagreement of subject and predicate, wrong case of pronoun, and similar grammatical errors. They conclude that Latin can be taught so as to be valuable in English, but that, as taught in the schools they investigated, it was more of a hindrance than a help in good English. These various investigations incidental to grammar indicate that, while Latin pupils are definitely superior to non-Latin pupils in grammar knowledge, this knowledge does not function in aiding the individual in his English.

The Plan of the Investigation

The present study was entered upon in an attempt to show what effect Latin is now having in a typical school situation. Specifically stated, the problem under investigation is this:

> Are Latin pupils, as they are taught in the typical secondary school of today, significantly superior to non-Latin pupils of equal mental ability in English vocabulary, spelling, and grammar?

The intention was to organize the study in such a way as to measure the effect of Latin on such factors in a situation typical of that in which the persons involved encounter English vocabulary, spelling, and grammar in common usage.

Six schools generously cooperated in the study by making available their school records, and in permitting the test designed for the study to be given to the seniors of their English classes. Four of these were coeducational high schools, namely, Edison and University High Schools of Minneapolis; Central High School, Aberdeen, South Dakota; and Chisholm High School, Chisholm, Minnesota. In addition, two Catholic academies, St. Joseph's Academy (girls), and St. Thomas Academy (boys), both of St. Paul, Minnesota, kindly consented to aid in the study.

These six schools vary quite widely, especially in that they include a wide range of social levels. The academies were included to give representation to private institutions at the secondary level. The schools are quite representative of schools of their size; however,

9. Coxe, Warren W., The Influence of Latin on the Spelling of English Words. Journal of Educational Research Monographs, No. 7, Bloomington, Illinois, Public School Publishing Company, 1924. Pp. 122.

10. Miller, George P. and Briggs, Thomas H., "The Effect of Latin Translations on English," School Review, XXXI (December, 1923), 756-62.

one may with reasonable assurance conclude that the teaching in these schools is superior to the general level of teaching in the average American high school.

Comparisons were made between the scores made on tests of vocabulary, grammar, and spelling as follows:

Unit I of the study:
Pupils having had two years of Latin vs. pupils having had no Latin—112 pairs.

Unit II of the study:
Pupils having had three or four years of Latin vs. pupils having had two years of Latin vs. pupils having had no Latin--29 trios.

In order to hold at a minimum the possiblity that differences discovered between Latin and non-Latin pupils might be attributed to differences in natural ability, pupils were so selected for study that the groups compared were approximated equal in ability. The pupils were paired on these bases--sex, chronological age, intelligence, economic status of family, the number of years spent in the study of modern foreign languages, and the first year English mark.

As an illustration of the equating of the groups, the data relative to the equality of the group which had studied two or more years of Latin, and those who had no Latin are given in Tables I and II.

The Tests in English Vocabulary, Spelling, and Grammar

The vocabulary test consisted of 38 multiple choice items, 19 of Latin origin and 19 of non-Latin origin. The words employed in this test were taken from the Inglis Test of English Vocabulary, the Purdue English Tests, and the Columbia Research Bureau English Test, most largely from the first named.

The spelling test was also of the multiple choice type and consisted of 40 items taken largely from the Seven S spelling scales, with a few from the Buckingham-Coxe Spelling Test and the Buckingham Extension of the Ayres Spelling Scale. One-half of the words were definitely of Latin origin and one-half were definitely not of Latin origin.

The grammar test included twenty English sentences, involving errors fairly common to high school classes.

As a measure of the reliability of the tests, coefficients of correlation were calculated between scores made on chance halves of the test and the coefficients of reliability of the entire tests, estimated by means of the Spearman-Brown prophecy formula. The resulting coefficients were:

Test I	Vocabulary	.70
Test II	Spelling	.65
Test III	Grammar	.50
The entire test		.78

The reliability of all of these was sufficiently great for reliable comparisons, based on the 112 pairs and for the 29 pairs, except in the case of the grammar test, when considered separately.

Relative Scores of Latin and Non-Latin Pupils

The relative scores of the Latin and the non-Latin pupils are given in Tables III, IV, V, and VI.

It will be noted that there are rather systematic differences in favor of those pupils who have studied Latin, and that on Latin words the pupils who had had more than two years of Latin excelled those who had spent but two years on the subject. The differences, with the exception of those between non-Latin and more than two year Latin pupils, are in general quite small. When compared with their probable errors, they are not great enough to be fully reliable.

With respect to the ability to spell, when the 112 pairs were considered, there was no significant difference between non-Latin pupils and those with 2

TABLE I

DATA RELATIVE TO EQUIVALENCE OF LATIN AND NON-LATIN GROUPS

	Mean Scores of Paired Pupils in Age and I.Q.						
		Age			Intelligence Quotient		
	N	Mean	$P.E._m$	S.D.	Mean	$P.E._m$	S.D.
Latins	112	17.59	0.04	0.61	107.50	0.70	11.00
Non-Latins	112	17.76	0.05	0.76	107.72	0.67	10.46

Average Marks in Numerals and Letters*

Kind of Mark	Latins	Non-Latins
Numerals (46 pairs)	87.41	86.99
Letters (66 pairs)	A, 12 B, 51 C, 62 D, 7	A, 7 B, 50 C, 67 D, 8

*The letters total twice the number of pupils in each school, since they are semester marks.

TABLE II

DATA RELATIVE TO EQUIVALENCE OF NON-LATIN, TWO-YEAR LATIN, AND MORE THAN TWO-YEAR LATIN PUPILS

Years of Latin		Mean Age and Intelligence Score					
		Age			Intelligence Quotient		
	N	Mean	$P.E._m$	S.D.	Mean	$P.E._m$	S.D.
Four-year Latins	29	17.62	0.06	.47	116.64	1.33	10.75
Two-year Latins	29	17.32	0.07	.58	116.64	1.51	12.04
Non-Latins	29	17.44	0.08	.64	116.29	1.37	10.95

Mean Number of Years of Study of Modern Foreign Languages

Four-year Latins	Two-year Latins	Non-Latins
0.66 year	0.67 year	0.63 year

TABLE III

MEAN SCORES IN THE VOCABULARY TEST

	Latin Words	Non-Latin Words	Total Test
Unit I			
Latin pupils	7.62 ± .14	9.99 ± .22	17.14 ± .32
Non-Latin pupils	7.59 ± .20	9.52 ± .22	16.55 ± .32
Difference	.03 ± .20	.47 ± .28	.59 ± .45
Unit II			
Four-year Latin pupils	8.59 ± .26	10.86 ± .43	19.93 ± .65
Two-year Latin pupils	7.38 ± .43	11.03 ± .43	19.31 ± .60
Non-Latin pupils	6.93 ± .65	9.83 ± .45	17.17 ± .81
Difference a-b	.76 ± .38	.17 ± .42	.62 ± .65
Difference a-c	1.66 ± .49	1.03 ± .53	2.76 ± .99
Difference b-c	1.45 ± .42	1.20 ± .50	2.14 ± .83

TABLE IV

THE MEAN SCORES IN THE SPELLING TEST

	Latin Words	Non-Latin Words	Total Test
Unit I (112 pairs)			
Latin pupils	16.96 ± .15	15.71 ± .14	32.25 ± .24
Non-Latin pupils	16.68 ± .16	16.20 ± .16	32.43 ± .29
Difference	.28 ± .19	-.49 ± .19	-0.18 ± .40
Unit II (29 trios)			
A) Four-year Latin pupils	17.69 ± .20	16.69 ± .25	34.83 ± .40
B) Two-year Latin pupils	17.38 ± .20	16.52 ± .30	34.34 ± .41
C) Non-Latin pupils	16.69 ± .33	16.10 ± .31	33.24 ± .70
Difference (A-B)	.31 ± .23	.17 ± .35	.49 ± .61
Difference (A-C)	1.00 ± .37	.59 ± .41	1.59 ± .80
Difference (B-C)	.69 ± .36	.42 ± .35	1.10 ± .72

years of Latin, the Latin pupils doing slightly better on words of Latin origin and the non-Latin pupils doing better on words not of Latin origin. When the 29 trios were considered, the 3-4-year Latin pupils were reliably superior on Latin words to the non-Latin pupils, but only slightly superior to the 2-year Latin pupils, and the 2-year Latin pupils were slightly superior to non-Latin pupils.

TABLE V

MEAN SCORES ON THE GRAMMAR TEST

Unit I (112 pairs)	
Latin pupils	15.80 ± .16
Non-Latin pupils	15.23 ± .16
Difference	.57 ± .24
Unit II (29 trios)	
A) Four-year Latin pupils	17.72 ± .23
B) Two-year Latin pupils	16.31 ± .30
C) Non-Latin pupils	15.59 ± .25
Difference A-B	1.41 ± .35
Difference A-C	2.13 ± .34
Difference B-C	.72 ± .35

With respect to ability to locate and correct the errors in sentences, Latin pupils are slightly though not reliably superior. In Unit II the two-year pupils were found to be reliably superior to the non-Latin pupils, and the four-year pupils superior to the two-year pupils, with an even greater superiority over non-Latin pupils.

As is to be expected from the preceding comparisons, the Latin pupils excel the non-Latin pupils on the entire test, especially on the Latin portions and particularly when the 3-4-year pupils were compared with the non-Latin pupils.

General Conclusions.- (1) Pupils who have had two years of Latin are likely to do slightly better in spelling, vocabulary, and grammer tests than pupils of similar natural general ability or intelligence who have not studied Latin. (2) Pupils who have had more than two years of Latin do materially better than those who have not studied Latin, and slightly better than those who have studied Latin for only two years. (3) The

TABLE VI

MEAN SCORES ON TOTAL TEST

Unit I (112 pairs)	Latin Portions	Non-Latin Portions	Total*
Latin pupils	39.30 ± .30	40.34 ± .38	63.80 ± .52
Non-Latin pupils	38.61 ± 38	39.93 ± .36	63.80 ± .58
Difference	.69 ± .39	.41 ± .44	.72 ± .65
Unit II (29 trios)			
A) Four-year pupils	43.55 ± .52	44.76 ± .73	71.41 ± .68
B) Two-year pupils	42.21 ± 5.6	44.55 ± .74	68.79 ± 9.3
C) Non-Latin pupils	39.86 ± .56	42.14 ± .70	66.17 ± 1.19
Difference A-B	1.34 ± 61	.21 ± .59	2.62 ± 1.13
Difference A-C	3.69 ± .93	2.62 ± .96	5.24 ± 1.56
Difference B-C	2.34 ± .84	2.41 ± .84	2.62 ± 136

*The total is not the sum of the non-Latin and the Latin portion, since in both of these classifications is included the entire grammar test which does not submit to division into Latin and non-Latin divisions.

differences noted above are not in general great enough to be statistically reliable with respect to words not of Latin origin. (4) The differences are so small as to invite suspicion concerning the degree to which Latin is so taught as to contribute most effectively to developing ability in English, spelling, vocabulary, and grammar.

ADDITIONAL BIBLIOGRAPHY

Gray, Mason D.,
Carr, Wilbur L., and
Grinstead, Wren Jones "Research in High School Latin" A Summary of the General Report of the Classical Investigation. Sixth Yearbook of the Department of Superintendence, pp. 372-382. Washington, D.C.: Department of Superintendence, 1928.

Robinson, Mary C. "An Experiment in Teaching Latin for the Sake of English," Classical Journal, XV (October, 1919), 42-49.

Thorndike, Edward L. "Gains Made in Ability in English by Pupils Who Do and Do Not Study Latin," School and Society, XVIII (December 8, 1923), 690.

Thorndike, Edward L. "The Influence of First-year Latin upon Range in English Vocabulary," School and Society, XVII (January 20, 1923), 82-84.

Thorndike, Edward L. "The Influence of First-year Latin upon Ability to Read English," School and Society, XVII (February 10, 1923), 165-68.

Wilcox, M. J. "Does the Study of High School Latin Improve High School English?" School and Society, VI (July 14, 1917), 58-60.

A GLOSSARY OF NEEDLESS READING ERRORS

by
F. L. Wells
Boston Psychopathic Hospital

Needless is here used in the sense that the child corrected the error (of oral reading) upon the mere knowledge that he had made some error. The children, thirteen in number, figuring in these observations are among those who were referred to the writer with at least partial reference to reading difficulties. In the laboratory exercises of oral reading which they had, record was kept of the errors according as they were (1) spontaneously corrected, with no participation of the examiner; (2) corrected upon merely being told that error was made; (3) corrected only upon more help than this. Only the first and second classes of these errors, those corrected without specific aid, figure in the present enumeration. In the range of the reading texts used, they markedly outnumber the third class; that is to say, the child fails more on words that he can read, than on words that he cannot.[1] The usual type of error is the second, where correction is not spontaneous, but follows on the mere indication that error has been made. This is shown in the listing as

horse = house

If, as is less common, the child corrects the error without being informed of it (the first class of error), the listing is with an asterisk,

*horse = house

Sometimes the listing bears a number, e.g.,

horse = house 3

This denotes that the child, told of error, still read the same horse as house three times altogether before correcting it. Again, he might on two or more different occasions read horse as house. In these cases the number of such occasions is put in parentheses, thus

horse = house (3)

This denotes that on three different occasions (repetitions elsewhere in the text or at different times) horse was read as house. Occasionally the child makes another and different misreading before correcting himself. These cases are listed

horse = house, home

This denotes that the child first read horse as house; upon being told he had made a mistake, read horse as home; again told he had made a mistake, read horse correctly.

Each entry in the listing is also preceded by a numeral. This denotes the individual making the error or errors of that entry. Thus,

4 horse = house (2)

denotes that Case 4 on two different occasions read horse as house. Of the thirteen cases, 2 and 4 are girls, the remainder boys. Seven of them figure in another study by Card and Wells, the corresponding data on age, grade, I.Q. and reading tests being repeated on the following page.

The reading material yielding these errors necessarily varies from child to child, as does the amount of reading involved. The most difficult is the text of Treasure Island; the simplest is a series of sentences including only

1. Responses that are produced under the third class (uncorrected without specific help) do not differ superficially from the other two classes. Instances from the record of Case 4 are: bath = beach; charge = cash; clearly = closely; friend = finest; nine or = nice of; politics = publicity; sled = seat; They = Nancy; this = that.

Case	1	2	3	4	5	6	7	8	9	10	11	12	13
Age at beginning observation	12-2	10-8	8-11	12-3	10-3	10-11	9-0	11-2	10-2	10-0	11-3	9-5	11-4
Grade placement observation	5-8	4-8	3-8	5-4	3-8	3-8	sp	4-3	sp	5-6	4-8	3-1	5-0
I.Q. score nearest observation	81	71	86	88	89	68	83	85	73	106	93	64	96
Grade Median of Reading Test Scores	3.4	3.1	1-	2.8	2.9	1.5	2.5	2.3	1.0	3.0	3.2	1.4	4.2

the commonest 60 words of the Gates Reading Vocabulary for the Primary Grades. Other material was based mainly on the first 500 words of the Gates list and on various school readers. A few word-lists also figure as above, partly from Thorndike or Gates, partly from Basic English. A word misread when thus out of context is preceded by the sign +. The printed response simply represents the vocable which was understood. Where this does not seem to be a word, it is put in quotes.

The source of these needless errors is to be considered from the two standpoints of why and how; why any error occurs, and how, given the conditions for error, it comes to take the form it does. The failure to read the word at the moment, in the presence of general ability to read it, denotes a phase of lowered integration. Popularly it is thought of as a phase of "inattention"; and if one etymologizes this word, one finds it in strict harmony with Janet's abaissement du niveau, or de la tension. This tensional regression may arise in a variety of ways; it is well known for example that reading difficulties are a way in which some children react to difficulties of home adjustment,[2] as others might respond with petty stealing, or with tantrums.

Such phases of tensional or "attentional" regression set the stage for errors like these. If there were not this regression, the error would not be made; and in several of these cases, if the child had misread only words he could not read, he would scarcely have appeared a reading problem. The verbal pattern then spoken in error, is the one whose threshold happens to be lowest at the moment. There are many factors tending towards a lowered threshold of this kind, not mutually exclusive, and perhaps never operating one alone. Thus if in a word-list, home is read as house, it is fair to assume that cues have come from both form and meaning. Where paws is read as feet, meaning has probably given the much more important cue, though the words are indeed the same length. The common error a = the is a function of meaning, particularly in its relation to context, rather than form. In connected reading matter it is not easy to separate the intrinsic meaning of a word from its contextual meaning; it is illustrated however in errors like that of drum = door. This occurred in a sentence "my ear is close to the drum." In this context, door is a much more natural word than drum, "its threshold is lower", and under tensional regression it gets the final common path. Similarly in a sentence "you can put the paint in the pan", pan is read as pot, with the additional cue of similar initial letters.[3] These cues of meaning and context have long been understood as factors in the sensory and motor lapses of language studied in adults

2. The errors here concerned are thus the sort that should first disappear under better motivation and improved general adjustment of the child, factors on which Durrell has put principal emphasis in the management of reading problems.
3. It may be remarked that this initial similarity and rough configurations such as length, appear the only formal cues of importance in this material.

by various observers, and figuring more recently in the Psychopathologie des Alltagslebens of the psychoanalyst; to this literature the interested reader may refer. The analyst has sought the cues mainly in repressed material; here they seem to be more superficial. Conventional meanings, both intrinsic and contextual, together with physical form, supply the essential cues. But cues may come from any possible associative source; in a word-list for example, engine is read as curtain, mainly because the latter word happens to occur elsewhere in the same word-list.

Another factor in the "why" of these errors, is the dislocation that exists, peculiarly in English, between the auditory and visual symbols of language. Better to understand this factor one needs comparable data from other languages such as German or Russian, whose spelling is much more phonetic, or French where it is less directly so, but spelling still offers a reliable cue to speaking. Not to say at the other extreme, the ideographic Chinese. As concerns English, it leads naturally enough to a scheme of teaching reading that minimizes the written word as a phonetic symbol, and teaches it as an ideograph (with resulting "underanalysis"). There is generally conceded to be a gain in reading facility, particularly speed, within the limits of reading as a teaching topic. But if the price paid is to read when one knows better, with a similar effect upon subsequent habits of thinking, the value of the gain is problematical.

The full tabulation follows:

7	+ a = an	4	a letter = his letters
8	a = an (2)	3	a mouse = the mice
3	a = and	5	about = above
6	a = and	7	about = after (3)
4	a = his	11	above = along
8	a = his	4	across = the roses
8	a = I	10	Admiral Byrd = a man blind
7	a = it	11	after = afterwards
13	a = on	2	after = often
1	a = the (4)	7	again = after
2	a = the (7)	1	*again = against 2
2	*a = the (3)	7	*again = around
4	a = the (12)	4	ahead = each
4	*a = the (2)	10	air = seen
5	a = the (10)	9	all = at 2
5	*a = the	5	all = best
6	a = the	4	all = called
7	a = the	1	all = old
7	*a = the	7	*all = on
8	a = the (4)	2	all = the
8	a = the 2	13	*all the little birdies = at the little babies
9	a = the	4	*alone = along
11	*a = the	5	also = lost
13	a = the (6)	7	always = alone
13	*a = the	5	+ am going = made, had
7	a = to	8	+ am = and
7	*a = to	1	an = a (2)
2	*a = we	3	an = and
1	a = you	2	an = in
4	a = you		
13	A box and paper = The box of paper		
8	a fan is good = a fish is got		
8	a hard = to have		

Admiral Byrd	as	a man blind	man	as	fellow
behind	"	himself	mice fell	"	nice fellows
Christ	"	Jesus	moccasins	"	Mississippi
Christmas	"	Santa Claus	nest	"	house
dress	"	cross	nose	"	mouse
ear	"	eye	on a stick	"	under a tree
eye	"	ear	paint	"	peanut
floor	"	ground	rat	"	mouse
good-bye	"	go away	squirrel	"	mosquito
hard	"	best	time to come over	"	something to eat
I	"	you	twenty	"	twelve
Jesus	"	Joseph	very soon	"	with some
kind	"	great	warm	"	winter
large	"	little	yard	"	garden

1 an = the
7 an = the
11 and = as
8 and = at
13 and = at
3 and = but
11 *and = but
7 and = dog
2 *and = of
7 and = had
11 and = his
1 and = I (2)
7 *and = my
9 *and = ran
1 and = the
4 and = the
5 and = the
7 and = the
13 and = the
1 and = to
4 and = to
8 and we saw = and waved so
2 another body = any other day
7 answer = animal
8 any = another
8 any = my
9 are = asked
4 are = is
13 are = were
2 are better = as butter
5 arm = band
3 arm = bands
7 arms = marches
8 around = under
7 around = up
8 around = up
2 *as = at
7 as = at 2
7 as = last
1 as before = as if before
11 as he had = and he said
4 as you = at your
7 +ask = case
8 ask = like
4 ask = speak
8 ask = think
4 *asked = said
3 at = a
7 at = an
8 at = an
4 at = and
7 at = ate
6 at = can
13 at = for
8 *at = it
13 at = not
5 at = that
13 at = that
4 at = the
7 at last = let's
3 ate = and
8 ate = ate up
6 ate = can
5 ate = eat
3 ate = is
4 away = always
2 *away = around

3 baby = boy
4 baby = boy
7 *baby = little (2)
7 baby = little 3
7 back = make
7 *bad = beds
7 bad = did
4 bad = dog
8 back = by
13 bag = bows
4 bag = dog
1 balancing = balanced
4 bank = side
8 bare = big (2)
8 barn = burn
4 +basket = biscuit
4 basket = biscuit
11 be = become
8 +be = bird
8 be = boy
4 be = buy
8 be = by
1 be = have
3 *bed = bear's
4 before = by (2)
7 beg = big
7 began = begged
4 *began = begin, begun
13 began = begun
1 beheld = behind
5 behind = himself
3 bell = box
7 bell = picture
2 best = beast
5 best = beat
8 best = boat
8 big = box
6 big = tree
13 big cats = a big cat
7 big clown = brown
7 bill = hill
1 Billy = Bill
3 *bird = bee
9 bird house = bird's house
4 birdies = babies
8 birdies = birds
3 *birdies = chickies
1 black = back
2 blankets = = baskets
7 blows = blew
6 bluebird = birdie
8 bluebird = butterfly
3 bluebird = butterfly, birthday
3 bluebird had = bluebirds has
4 boat = boot
4 *boat = boot
4 +book = box
4 boot = boat (2)
4 +boot = boat (3)
7 Boston = bushes
8 both = booth
3 bowl = bag
6 bowl = basin
2 boy = baby
7 *boy = baby
4 boys = boy
8 boy = boys
12 *boy = girl
8 +bread = bird
8 bring = get (2)
13 broom = bloom
1 brown = brow (2)
13 but = and
7 but = be
8 but = by
8 *but = by (2)
11 but he = by his
8 but I = by the
5 but = put
6 buy a = bring the
8 bull's = bull
11 by = but
1 by = with

6 cake = candy, can
7 cake = care
8 cake = catch
4 came = can
7 came = can
8 *came = can
7 came = come
11 +came = come
8 came = comes (3)
13 came = comes
8 *can = came
12 can = come
12 *can = come
1 can = could (2)
13 can blow = blew
9 can eat = cannot
4 can fly = came flying
4 can't = can (2)
1 can't = cane
13 can talk = came from
13 can you = you can
4 cap = "cate"
5 cap = hat
6 car = auto
4 carry = cry
8+*cart = coat
4 castle = camel
4 castle = palace
4 cat = coats
7 cat = dog
5 cat = kitten
8 catch = chase
6 chain = child
1 change = chance
4 chased chained (2)
13 chased = chases
5 chick = chicken (2)
5 *chick = chicken
8 *chick = chicken
7 chicken = hen
7 chickens = children
5 child = children
8 child = children
6 child = chin
3 children = a child
8 chin = catch
8 +chin = skin (2)
4 Christ = Christmas
4 Christ = Jesus
3 Christmas = Santa Claus
3 city = child
1 clear = cheer
5 clear = clean
7 +climbed = chained
5 climbed = count
5 climb up = jump over, jump up, climb over
5 close = covered
13 close to = closed by
8 clothes = color
4 +clouds = cloud
7 clouds = crumbs
7 *clown = circus
8 *collar = color
7 cold = old
8 color = call
6 color = cat
5 color = cool
6 color = corn
7 color = draw
4 come = came
9 come = came
4 *come = can
8 come = can (2)
12 come = can
7 *come = comb
2 come = comes (2)
7 come = some
4 coming = calling

1 complaint = complaining
8 *cook = book
6 corn = curtain
3 corn = grass
6 corn = hay
7 could = can 2
4 could = couldn't
4 could = cried
7 could = need
7 could = round
8 +course = cross
11 cover = clothes
3 cradle = crib
6 crayon = coat
13 crayon = crown
4 cries = tears
8 cross = close
4 cross = rose
6 cry = city
8 curtain = camera
8 +curtain = cute
8 curtain = engine
1 cutlasses = cutlass

8 Daddy = Dad
5 dancer = dainty
4 *danger = dinner
4 *dark = deep
8 +dark = duck
4 dearest = destitute
1 *death = dead
1 did = do
7 did = do
9 did = do
4 did = does
5 did = said, could
4 *didn't = don't
7 dig = be
7 dinner = dead
4 dinner = sinner
8 *Dish = drink
7 dish = shut
1 dish with = dash which
4 dish with = ditch which
8 do = can
8 do = can 2
8 do = did
8 *do = did
4 Do = Did you
8 do = don't
8 do not = don't
2 do you = you do
4 does = did
5 does = got
4 does not = don't
8 +dog = bird
7 doing = drawing
5 doing = going
5 doing = going to
4 doll = collar
2 doll = dolly
8 *doll = dolly
7 donkeys = monkeys
7 don't = do not
8 *don't = do not
9 door = poor
7 down = now
7 dress = cross
4 *dress = dishes
4 dressed = dresses
5 dressed = dresses
5 *dressed = dresses
4 *Dresser = Doris
4 drop = drip
4 dropped = dripping
4 dropped = dropping
2 *drum = door
8 dry = weather

4 each = reached
8 *ear = eye
2 ear = eyes
7 ear = read
1 earlier = easy
7 ears = eyes (2)
8 +east = earth
4 eat = sit
4 eight = enough
7 end = red
3 enemies = "enemies"
4 +engine = curtain
4 equal = make (2)
4 *even = every
11 ever = even
2 ever = every
5 ever = every (2)
7 ever = every (2)
11 ever = every
4 ever = never
11 +eye = ear
4 faithfully = friendly
4 *fall = fill
7 far = fat
4 farm = fair
10 farm = fir
8 farm = front
5 farmer = farm
4 *fast = first
2 fastened = finished
4 feast = palace
5 feel = fell
8 feet = foot
7 fell = full
13 fellow = fellows
4 felt = left (3)
4 field = fleet
4 field = sheet
4 fifty = five
4 filled = "fulled"
9 find = flew
7 find = found
8 +fine = fin (2)
4 fine = fire
8 +fine = fire
7 fine = for
8 +fine = line
4 fine = tiny
13 find = found
5 fire = fence (2)
13 *fish = dish
4 flat = floor
7 flew = "flied"
8 *floor = ground
13 floors = flood
4 flower = flowers
4 *fly = find
8 fly = fill
5 fly = flew
8 fly = flew (2)
9 fly = flew 2
13 folded = followed
2 *foot = feet
4 foot = feet
5 foot = feet (2)
7 foot = feet
13 foot = felt
7 foot = food
4 *foot = front
5 for = from
7 for = from
7 *for = fun
3 for = if
6 for = is
2 for = of
3 for = of
7 for = of
3 for = of, if
5 for = to
8 *for = two
1 former = form
11 fought = found
9 found = flew
7 found = forced
2 found = from
4 *found = saw
8 four = far
2 four = foot
11 +four = fourth
7 four = run
4 fox = box
4 fragrance = fragrant
4 friend = family
4 friend = fried
7 friendly = friends
4 friends = father
7 friends = fathers
4 from = for
7 form = for (3)
8 from = for
2 from = of
8 from = of
4 from = off
5 from = with
1 frosts = "frozes"
5 full = felt
4 fun = fine
8 funny = full of the

4 gave = going
2 *gave = got
8 gave = got
4 gave = give
13 gave = grave
13 gave = grove
1 get = go
7 get = tell
3 girl = good
8 +give = gave
7 give = get
9 give = go
8 give me = get my
4 given = gave
11 +given = give
8 +glass = glove
4 glove = gloves
8 go = good
8 go = got
8 go to = got out
7 goat = gate
5 goes = go
5 *going = doing
2 gold = golden
8 +good = God
8 good = great (2)
7 *Good-bye = Go away
2 got = get
1 got his = gone to
4 grass = glass
4 grass = green
4 grass = guest
5 grand = grouchy
4 grandma = mother
2 grandmother = grandma (2)
4 grandmother = grandma
5 grandmother = grandma 2
5 grandmother = grandma
8 *grandmother = grandma (2)
7 greedy = great (2)
7 +greedy = greatly
7 grew = came
1 grip = grasp
7 ground = garden
7 growl = grow
5 grows = goods
4 grows = grew

7 had = and
2 had = has (2)
3 had = has
8 had = has (2)
12 had = has
2 had = hard
9 had = he said
4 had = made
5 had = made
7 had = said
8 +hair = head
7 happened = helped
8 +hand = head
5 hands = arm
8 hand and we = hand and waved
4 happens = happened
4 hard = best
4 hard = hardly had
13 has = can
2 *has = had
4 has = had (3)
5 has = had (3)
11 has = had 2
12 has = had (2)
2 has = has, have
2 has = have
8 has = his
11 +has = his (2)
8 has = is (2)
13 has = is
13 *has = is
7 has = was
6 has = you, hers, his
4 has no = has not
4 has no = was not
4 has no feet = was not fed, has not feet
4 has some = was had
5 have = be
8 have = every
8 have = gave
2 have = had
4 have = had
5 have = had (3)
4 *have = hear
4 have = never
4 have = take
4 have seen = had since
8 *have you seen = here you can
11 he = and
8 he = here (3)
5 he = I, we
1 he = it
13 *he = it
4 he = him (2)
11 he = how
4 he = she
7 he = the (4)
7 he = there
5 he = we
7 he = we
5 he gave = we came
4 *he held = we need, hemmed
13 *he held = we have
1 he kept = to keep
8 +head = hair
4 head = hand
7 head = hand (2)
13 head = hand
7 hear = have
7 heard = had, hid
4 heaven = nearer
4 heavier = heavy
7 +helped = happy
8 helps = held
13 helps to take care = care to take
4 *her = here
7 her = here
8 her = here (2)
8 *her = here
11 +her = here (2)
5 her = his
4 her = she
7 her = the (2)
4 her hand = the land
1 here = he
5 here = her
8 here = how
9 here = how
4 here = there, those
7 here = this
7 here = where
13 *here = where
7 herd = hat
7 high = big
7 high = home
2 him = her
4 him = her
12 him = his
4 him = me
5 him = them
2 him on the nose = her on the head
7 hind = head
8 his = I
7 his = he
6 his = here
1 his = the (2)
7 his = the 2
7 his = the
8 his = the
4 his = this
4 *his = this
7 his = this
4 his head = this hand
1 his house = house here
5 hit = put
7 hold = told
4 holy = hail
2 *home = house
4 home = house
7 home = house (2)
11 home = house
2 homes = houses
2 *homes = houses
5 hop = hopping
11 +hope = hop
8 hope = hot
4 horse = house
4 +horse = house
11 horse = house, home
2 *house = home
4 *house = home
9 *house = home 2
13 *house = home
4 house = horse
4 *house = horse
5 house = horse
8 +house = horse
1 how = now
4 *how = now
4 how = now
9 hurt = has

1 I = and
8 I = he
1 I = I'll
1 I = I've
4 I = it
2 I = of
2 I = the
2 *I = the
11 I = to
8 I = you
8 I = we
4 I be = I've been
4 I'll = I've been
5 I'll = I will
4 icy = ice
13 if = or
3 in = and
3 *in = at
9 in = he
4 *in = I (2)
8 *in = into
8 in = is (4)
1 in = it
2 *in = it
7 *in = it
9 in = it
7 in = on (2)
8 in = on (2)
13 in = on
13 in = on the
5 in = the
7 in = to (2)
9 in = us
7 in his = this
4 in his = with this
13 in summer = in a sweater
4 indeed = in need
4 instead = instantly
13 *is = all
13 is = are
13 *is = does
13 is = has (2)
9 is = he
9 is = here
8 is = in
13 is = its
13 is = of
1 is = this
1 is = to
4 is = was
13 is = was (2)
13 is looking = looks
7 is one = some
7 is the = it is
1 Is the = this is
4 *is this = it is
7 it = he
13 it = him
8 It = I
2 it = is
4 it = this
4 its = this
7 it = to
7 its = the

4 Jesus = Joseph
11 John = Johnny
5 jolly = joy

2 keep = kept
8 Kendall = Canada
1 kept = keep
1 kept = keeping
7 kind = great (4)
4 *kite = kitty
7 kitten = cat (2)
4 *kitten = kitchen
5 kitty = cat
7 knocked = broke
13 *know = knot
11 known = know
8 knows = looks

8 lady = land
13 *laid = led
5 lake = lady
4 large = larger
8 large = little

1 laying = lying
7 learned = ready
4 learns = learning
2 leaves = leaf
5 leaves = left
4 leaves = love
13 Ledge = Lodge
5 leg = left
4 lessons = les-son
1 lest = unless
9 let = told
4 letter = little (2)
7 lifted = pulled
8 like = long
8 like = look
8 likes = like
8 like this = likes to
7 liked = lied
8 likely = like
7 +limb = climb
7 limb = round
7 *little = baby
5 live = lay
1 live = laugh
8 live = like
7 lived = lied
5 lived = loved
13 *lofty = "lofly"
7 lonely = alone
8 long = live
5 look = like
4 look = looks
8 look = looks
5 looks = looked
6 looked = caught
6 looked = child
4 looked = looks
5 looked = looks (2)
13 looked = looks
9 looked = "tooked"
6 looked at = liked to
4 looked the second = liked the sound
4 looser = last
7 lost = long
11 loudly = loud
4 love = lived
4 loved = lived (2)
4 loved and = lived the
1 lump = lamp
1 lying = laying
11 lying = laying
11 lying = leading
5 made = can make
5 made = have
2 made = makes
4 made = named
13 make = like 2
2 make = made
4 make = take
13 makes = making
4 mama = ma
8 man = fellow
11 man = name
8 man's = man
5 many = made
2 *may = is
4 me = a
4 means = names
2 member = number
2 men = man
4 mending = knitting
4 mice = nice
13 *mice fell = nice fellows
5 miles = middle
8 Miss = Mister (2)
13 Miss = Mister
4 Miss = Mrs. (2)
5 Miss = Mrs. (3)
8 Miss = Mrs. (3)
8+*Miss = Mrs.
7 Mistress = Master
2 moccasins = Mississippi
8 more = men
8 more = minute
7 more = mother's
8 more = moment
8 most = almost
8 most = next
7 +most = some
2 most often = must open 2
4 mouse = horse (2)
8 mouse = mice
4 mouth = month
2 mouth = nose
7 moving = circus
7 *Mr. = the
13 Mrs. = Miss
4 much = which
1 mutilated = muddy
3 my = and, by
11 my = may
4 *name = man
8 name = man
7 near = neat
7 near = neither
9 near = now 2
9 near = now
7 near = there (2)
4 nearly = near
1 neighbors = neighborhood
7 nest = house
4 nest = meat (3)
5 *never = any
4 nine = meat
7 no = on
1 no = not
9 no = not
9 no = not 2
2 nose = head
8 nose = head
8 nose = mouse
8 Norway = Norwegian
11 not = do
7 not = don't
5 not = hot
2 not = no
9 *not = no
9 not = now
5 not = you
4 now = how (2)
2 *now = not
7 now = on (2)
11 +now = no
13 now = then
2 number = members
11 number = members
7 number = under
5 nut = nest
5 nut cracker = cocoanut
5 oak = oaktree (2)
4 of = at
4 of = for (4)
6 of = for
7 of = for (4)
7 of = for 3
8 of = for
9 of = for
9 *of = for
11 +of = for
5 of = from (2)
9 of = from 2
2 *of = in
13 of = in
8 of = on
4 *of = to
13 *of = to
5 of = who
4 of course = the corner
8 of our = for your
11 off = for (2)
11 +off = for
8 +off = of
7 off = over
8 off = over
7 off = up
2 often = after
4 often = after
4 often = have to
9 *Oh = us 2
13 older = other
7 *on = and
8 on = around
4 *on = in
8 on = in
13 *on = in (2)
13 on = in
7 *on = in
1 on = of
7 on = no (2)
8 on = no
9 on = no
8 on a stick = under a tree
13 on spectacles = spectacles on
2 on the ship = the sea
7 one = day
1 one = on
5 one = the
7 ones = nose
4 one dark = on the back
13 one of his daughters = his daughter
7 only = your
2 or = of
5 or = of
13 or = of the
5 or = nor
5 or = you
11 +order = other
13 other = others
4 our = you
4 our = your
11 our home = your house
9 out = down
4 out = on
7 out = up
13 out to walk = out on a walk
7 *over = every
7 over = every
8 over = every
4 over = our (2)
4 *own = one
4 own = only
4 owned = only
8 pack = bag
4 paint = peanut
2 paint = plant
4 paint = plant
4 pan = plan
5 pan = pot
3 paper = apple
8 *paper = apple
4 paper = parlor
6 paper = "play apple"
7 park = picture
7 part = party 2
2 passed = passing
8 pasture = picture 2

4 paw = pot
7 paws = feet
8 pay = play
1 paying = playing
1 *peace = place
13 people = children 3
1 perfectly = painful
4 picnic = picture
8 picture = people
8 picture = plant
5 pictures = picture
8 pictures = picture (2)
4 pie = pig
8 *pig = big
8 pig = girl
8 pig = pigs
4 pig = pipe
5 pig's = big pig's
8 pipe = pie
4 +pipe = piece (2)
4 pity = put
4 *place = palace
8 place = "plake" (2)
2 plains = plants
13 plainly = plainty, planty
2 plant = paint
8 *plant = paint
7 plant = plate
7 +plant = plate
7 planted = played 2
7 plants = places
8 *play = stay
9 poor = door
4 pray = flag
4 pray = prayers
4 prayers = prayer
9 pretty = happy
8 pretty = pleased
5 pull = push
2 *pull = put
7 pulled = put
7 puppy = pussy 2
7 pushed = put
5 put = but
7 put = but (2)
5 put = buy
3 put = plays (2)
2 put = puts
8 put = put on
4 putting his hand = pulling his head

7 *rabbit = squirrel 2
7 rabbits = squirrels
11 racing = races
11 ran = run
5 rat = mouse
4 rather = never
8 raw = run
2 raw = runs
13 reach = reaches
4 reached = wished
2 real = regular
3 red = brown
12 red = sweating
12 red = tire
1 repent = repeat
1 replies = repeats
5 ribbon = robin
5 ride = road
4 ride = side
7 ride = write
4 right = left
7 right = night
4 +ring = rain (3)
4 ring = rain
1 risen = raised
7 river = rain
8 road = farm
8 road = fork
8 road = ran
8 *road = ring
4 roared = reared
4 robe = hot
11 Robin = Robin Hood 2
9 robins = bird
4 rosary = rosaries
4 *rose = cross
4 rose = reached
8 *roll = run
7 room = row
4 row = rowing
4 rubbing = rapidly
4 rubbing = rapping
4 rubbing = wiping
7 rug = room
12 run = ran (2)
13 ruts = roots (2)
13 *ruts = runts

4 *sadly = suddenly
7 safety = faster
7 +safety = softly
1 said = and
4 said = and
7 *said = and (2)
7 said = and
9 said = and
4 said = asked
4 said = says
4 sail = (the thing in a boat)
3 sat = is, it
6 sat = is sitting
7 saved = stopped
4 saving = serving
7 saw = said (2)
7 saw = say
2 *saw = was
2 saw = was (2)
4 saw = was (4)
5 saw = was
1 say = said
4 say = said
4 say = she
1 says = said
7 says = sells
3 saw = want
5 sea = sand
1 seaman = seafaring man
8 seat = set
8 see = say
8 see = use
4 seed = sand
5 seed = seeds
8 seed = seeds
4 seemed = second (2)
1 sees = seen
1 sees = seen, sun
4 seized = "frozed"
4 sell = send
4 servant = sewer
4 serve = sent
8 set = seat
4 set = seated
4 shaking = speaking
4 shall = should
4 *she = was
7 she = shall
7 she = so
4 +sheep = ship
8 sheep = ship (2)
8 *sheep = ship
4 sheet = sheep
2 sheet = short
7 shells = smell
2 *ship = boat
5 ship = side, still
1 shipmate = seamate
8 shirt = fish
13 shirt = short
4 shirt = skirt
4 +shirt = skirt
4 shoe = foot
4 shoe = shoes
4 shoe = snow (2)
11 shook = show
11 shook = took
11 +shore = sheet
7 should = shows
4 should rather = shall never
7 show = see
4 show = share
7 *show = sheet 3
4 show = should
4 show = snow
1 showed = slowed (2)
1 since = silence
5 single = smile
4 sins = son
4 Sisters = Saints
4 sisters = sister
13 sit = sits 2
8 sitting = sleeping
4 skirt = shirt
1 slammed = slamming
1 slamming = slammed
8 *sled = sail
10 sleds = sleighs
7 slippers = sleeper
8 smaller = small (2)
13 *smells = swells
11 smiled = small
4 smiled = smiling
4 snow = show
7 snow = winter
2 so = some
8 +soak = ship
4 +sock = soak
7 soft = some
5 soft = strong
5 soft = stronger
9 some = come
4 some = more
2 *some = same
9 some = soon
7 somebody = some people
7 somebody = some person
7 somebody = something
5 somersets = summertime
1 sometime = something
1 something = sometimes
2 something = sometimes
1 sort = short
2 south = southern
1 speak = spoken
4 spoke = stood

8 spots = "sheeps"
4 squirrel = mosquito
4 star = sister
8 star = stair
5 stars = star, street
4 stars = stairs
4 start = sheet
11 started = stood
4 steam = storm
8 steep = step
8 +step = steep (2)
8 step = steep
4 stick = switch
8 +still = style
1 stopped = stepped
8 story = city, store
11 +story = storm
9 straw = twigs
8 street = store
11 street = strong
7 study = story
8 sun = summer
11 +sure = such
4 surely = certainly
1 surely = scarcely
4 sweetest = sweet
1 swing = slash
8 table = bottle
13 *tail = tall
6 take = can
4 take = let
4 take = make
4 take = take, took
7 taken = snapped
13 taking = taken
13 taking it = talking
13 taking it easy = trying to escape
7 talk = speak
7 tar = mud 2
4 *task = looked
4 tell = take
1 tell = teach
7 tells = told
8 team = truck
8 ten = eight
2 tepees = trees
8 than = that
7 thank = that
9 thank = ask
4 that = as
4 that = at
5 that = at (3)
7 that = than
1 that = the (3)
2 that = the
4 that = the (2)
7 that = the
13 that = the
2 that = this
7 that = this
9 that = this
11 +that = thus
5 that = what
7 that = what (2)
7 that = what 2
7 that way = went away
2 the = a (2)
3 the = a
4 the = a (3)
4 *the = a
5 the = a 6
5 *the = a
5 the = a 2
6 the = a
7 +the = a
8 the = a (2)
13 the = a (2)
1 the = and
2 the = and
5 the = and (2)
7 *the = he
11 the = he
13 the = her
5 the = his (2)
7 the = his
7 *the = his
8 the = his
8 the = I
7 *the = in
7 the = it (2)
2 the = my
4 the = my
5 the = my (2)
8 the = my (2)
2 the = some
1 the = that
4 the = that
5 the = that
5 the = then
4 the = these
4 the = they
8 the = they (5)
2 *the = this
5 the = to (2)
7 *the = to
5 *the = you (2)
13 the = your
7 the = when 2
8 the little = the other little
8 the water = they cannot
4 *their = they
7 their = your
5 them = him
8 them = the
4 them = their
4 them = us
7 then = they 2
9 *then = they
4 then = when, where
5 then = when
7 then = when (5)
13 then = when (2)
1 then followed = the following
7 there = here
8 *there = here
10 there = the
11 +there = their
9 there = then (2)
11 there = these
5 there = they (6)
7 there = they
5 there = they, here
5 there = they were, where, they, there
8 there = third
8 these = this (2)
2 there = those (2)
11 there = those
4 there = three
8 there = three (2)
7 there = when
7 +these = this
8 these = this (2)
8 +these = this (3)
2 *these = those (5)
4 these = those (4)
10 these = those
11 +these = those (3)
8 these = three
8 they = the
11 they = their
2 they = their, that, this
7 they = then 2
7 they = then (2)
8 they = there (2)
13 *they = there
7 they = when (2)
12 they = you
1 they're = that you're
8 they were = there was (2)
9 thick = back
7 things = this
5 think we = drink we
5 thinnest = nicest
11 +third = these
8 third = three
4 this = a
13 This = A
4 this = it 3
1 this = his (2)
5 this = his (5)
7 this = his
8 this = his
7 this = side
3 this = that
4 this = that (2)
2 this = the (2)
3 this = the
5 this = the
8 this = the (2)
7 this = these
8 this = these
2 this = they
2 this = they, the
8 this picture = these pictures (2)
1 threatened = threatening
8 +three = the
8 three = there
7 +three = there
11 +three = these
8 three = third
7 through = there
1 through = thought
4 thrown = terrible
4 tie = tail
13 time = times
12 time to come over = something to eat
7 to = a
2 *to = and
5 *to = and
4 to = for
3 to = in, the
7 to = it
7 *to = in
7 to = it
7 to = it 4
7 *to = it (4)
7 to do = it did
7 to = it
6 to = the
7 *to = the
7 to = the
8 to = the
1 to = you
8 to do this = today
5 to eat = for food
2 to fly = and flew

13 today = one day
7 together = taking
11 told = took
1 took = got to take
8 took = take (4)
9 took = look (3)
4 *took = looked
9 took = looked
7 took = told
7 tops = big
4 toy = top
5 toy = top (2)
8 +tree = table
4 tree = three
5 trees = green
2 *trimming = tribe
4 *turn = trade
9 twigs = dirt
9 twigs = straw
5 twenty = ten (2)
2 twenty = twelve (2)
8 under = around
5 under = on
7 under = until
8 under = up
2 under = used
7 until = under (2)
7 *until = under
7 until = uncle
13 us = her
8 us = is
12 *us = them
2 *us = up
4 using = used

8 very = ever
7 very = every
7 very = with
9 very soon = with some
4 Vogel = goal
4 Vogel = Govell (3)
4 Vogel = Govel, Voger

11 *waded = went
7 wagged = wagon
7 wagon = cart
7 wagon = walking
8 wait = wet
2 *walk = water
8 wall = table
8 *wall = was
2 wall = walk
7 want = must
7 want = went
8 want = went
1 wanted = wants
7 wanted = went
7 wanted = went 2
7 wanted = would
1 wanting = waiting
5 warm = wild
9 warm = winter
4 was = as
13 *was = has
2 was = is
4 was = is (2)
7 was = is
8 was = is
13 was = is (2)
8 was = saw (2)
7 *was = walked
9 *was = went
9 was = went (2)
1 was = went
5 was = were
11 was = were
2 was called = you call
8 *was sitting = sat
13 was taken = is taking
8 wash = wet (2)
9 washed = went
4 watch = water
7 watch = witch
9 *watched = looked
13 *water = want
6 water = was, where
4 water = wind
10 waving = flapping
10 waving = wagging
1 we = he
2 we = he
5 we = he (2)
7 we = he
2 we = you
4 we = you
8 we = you 3
1 we = we'll
7 went = met
9 went = saw
4 *went = want
4 went = want (2)
7 went = wanted (2)
7 went = was
9 went = was
9 went = were
13 were = are
1 were = was
2 were = was
4 were = was
5 were = was
7 were = was (4)
9 were = was
5 *were = there
4 were = there was
5 were = they
8 were = went
8 were = you
4 we're = we've
4 we've = I've
8 what = at
8 *what = at
4 what = that
7 what = that
4 what = when
7 what = when (2)
9 what = when
9 what = where
1 what = well
1 what a = as I
4 what kind = that the King
13 *when = for
4 when = then (3)
7 when = then (3)
7 when = then 8
7 when = they
4 *when = where
4 when = where
4 when = why don't
6 when evening = whenever, when every
1 when the = with he
8 when you = we can
13 where = there
7 where = we
8 +where = were
11 +where = were (2)
4 where = who (2)
5 where were = they, they
4 which = such
11 +which = watch 2
8 +which = wish
7 which = with
5 which looked = such looking
7 white = little
4 who = how
4 who will = how well
4 whom = when
4 why = what (2)
2 wide = wild
5 will = can
7 will = can
4 will = while
7 will = with
8 will cry = was crying
4 wind = wide
4 wind = wound
8 *window = windows
7 *winter = snow
4 winter = warm, water
1 with = and
7 +with = these
7 with = they
1 with = was
1 with = when
4 with = when
7 with = when
9 with = when
4 with = which
13 with = which
1 with = we'll, will
4 with new = they now
13 without = for
4 wish = witch
4 wish = worship
2 woman = women
8 world = work
13 wrist = waist
6 writes = was
7 wrote = worked
7 would = wanted
7 would = was (2)
4 would = could (3)
7 would = could
4 would = couldn't
7 would = round
1 would = will
8 would = will
4 would = will, should
1 wouldn't = would (2)

7 yard = garden
1 you = a
11 you = be
4 you = for
11 you = he
7 you = him
2 *you = I
8 you = I
11 You = I'll
4 you = one
3 you = the
4 you = the (2)
4 *you = they
4 you = to
3 you = your (4)
8 you = your (2)
8 +you = your
8 you = we (3)
8 *you = we
8 you = we 3
1 you = who
4 *you both = your teeth
1 you may = who many
5 you would = we could
7 your = their
8 *your = you
8 your = you
12 Your = You (2)

AN EXPERIMENTAL INVESTIGATION OF THE EFFECT OF METHOD OF PRESENTATION UPON THE IMAGINATIVE QUALITY OF DESCRIPTIVE WRITING AMONG ELEMENTARY-SCHOOL PUPILS[1]

by
Maxwell F. Littwin, Principal
Forest Park Public School
New York City

I. The Problem and Its Importance

This investigation was undertaken in the hope that it might be possible to evolve definite methodology by means of which imagination in the descriptive writing of elementary-school pupils could be most effectively stimulated. The increasing interest in creative youth and creative writing has brought the teaching of composition more than ever to the fore. As Professor Hughes Mearns[2] expresses it, creative writing is a challerge to the drudgery of set class assignments by the teacher, the conventional theme, and the belief that the absence of mechanical errors is the chief aim in composition teaching.

The problem of teaching composition to children is at the present time a controversy between traditional methods and the methods suggested by creative education and progressive educators here and abroad. It is because of a realization that the existing interest in creative composition ought to be given greater opportunity for expression in teaching methodology that this research was undertaken. An experimental evaluation of the effect of different methods of presentation upon the imaginative quality of pupils' descriptive writing may be helpful in throwing light upon a phase of pedagogy that is still much shrouded in darkness.

II. Historical Background of the Problem

For some time traditional motions regarding the aims and purposes of composition teaching have been coming in for persistent attack. Attention is being gradually diverted from the composition product to the composition process, from the number and variety of formal and mechanical errors made by pupils to the character of the observation, imagination, and reflection shown by them. Attempts to measure "general merit" in the composition product have been shown to be futile. Composition is a highly complex process, possessing many "intrinsic" qualities, each of which can be studied effectively only when isolated, or when all other factors are equalized, or where they are known and accounted for. The trend in the measurement of language has been from a gross measure of general merit to a specific measure of individual factors which are included in this general merit.[3] Not only is the present investigation, therefore, timely in the sense that it is in line with the growing interest in creative writing; but in attempting to evaluate a specific element in composition--the quality called

1. Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education of New York University, 1934.
2. Mearns, Hughes, Creative Youth (New York: Doubleday Doran, 1929), 129.
3. Ashbaugh, E. J., "The Measurement of Language," Journal of Educational Research, IV (1921), 59.

imagination--instead of general merit, it is in accord with the present tendency in the measurement of language.

III. Previous Investigations

Previous investigations in the field of composition teaching shed only an occasional glimmer of light upon the problem involved in this study. Most of these investigations are concerned with subjects that are of greater interest in the field of psychology than in the field of pedagogy. Studies designed to test the capacity of pupils to attend and observe;[4] to determine to what extent children observe and how much they observe;[5] to evaluate the effect of time interval upon the fullness and fidelity of report;[6] to study in a concrete way the imaginative processes in a number of normal youths;[7] to investigate the validity of various methods used to determine ideational types;[8]--all these are of more concern to the student of psychology than to the student of teaching method. The investigations that have a closer relation to the present problem are Colvin's inquiry[9] as to whether it is possible to stimulate originality and spontaneity in pupils' compositions and Colvin and Meyer's study[10] to discover the imaginative elements in the written work of school children.

Furthermore, most of the investigations mentioned are to a large extent lacking in scientific technique. They are usually only preliminary tests and need to be supplemented by real and control tests. The investigators, to judge by the reports, seldom took into account the internal and external conditions involved in the experimentation; nor did they keep the conditions uniform and controlled or vary them purposefully in order to isolate the influence of these conditions--a most important element in scientific research.

IV. Experimental Procedure

Working Hypothesis.- The experimenter took as his working hypothesis the following: A well-planned, definite teaching procedure in a certain subject of the school curriculum, if extended over a period of time, will result in a measurable amount of pupil improvement in that subject. Furthermore, if two or more methods of teaching a particular subject of the school curriculum are employed over a period of time, with all factors except the teaching method remaining constant, an accurate measure of pupil achievement at the end of the teaching period will be a fair measure of the relative effectiveness of the teaching methods used.

Schools and Pupils Involved.- Four schools and a total of 961 seventh- and eighth-grade elementary-school pupils, ranging in age from 11 years, 0 months, to 16 years, 10 months, were involved in the experimentation. Two of the schools, Experimental School A and Control School C, were public schools in the Queens section of New York City; they are referred to in the discussion as the Alpha Environment. The other two schools, Experimental School B and Control School D, were public schools in Brooklyn. These are referred to as the Beta Environment. All together, 32 classes were included in the study, three for each 7A, 7B, 8A, and 8B grade in the experimental schools, and

4. Whipple, G. M., Manual of Mental and Physical Tests (Baltimore: Warwick and York, 1910).
5. Winch, W. H., Children's Perceptions (Baltimore: Warwick and York, 1914).
6. Slattery, E. S., "Effect of Time Interval, etc.," doctor's thesis, New York University, 1912.
7. Brittain, H. R., "A Study in Imagination," Pedagogical Seminary, XIV (1907), 137-207.
8. Wharton, W. P., "Experimental Study of Ideational Types," doctor's thesis, New York University, 1911.
9. Colvin, S. S. "Invention vs. Form in English Composition," Pedagogical Seminary, IX (1902), 393-421.
10. Colvin, S. S. and Meyer, I. F. "Imaginative Elements in the Written Work of School Children," Pedagogical Seminary, XIII (1906), 84-93.

one for each of these grades in the control schools. The investigator hoped that in these groups he had secured a "sampling" that was fairly representative of the cosmopolitan character of the New York City public-school population. The experimentation was undertaken with pupils of the seventh and eighth elementary grades, because such pupils are of an age when the emotions and the imagination, having become increasingly active,[11] should be expected to be reflected in their written work.

Equating the Groups.- In order to have approximately equivalent groups upon which to try out the effectiveness of the teaching procedures, the investigator classified the pupils on the basis of their intelligence quotients. In the Alpha group of schools, this classification was carried out not only for the entire group using a particular teaching method, but for each grade as well; for the Beta schools, this classification was possible for each total teaching group, but not for the various grade groups within the total.

Procuring the Data.- In addition to preliminary data on the socio-economic and racial-national background of the pupils involved in the study, the experimental data were secured on the basis of two composition tests given to the participating subjects--an initial test, at the beginning of the ten-week training period and a final test, given at the conclusion of the training period. The two tests were on subjects within the comprehension and interest of the pupils, as well as being easily susceptible to imaginative treatment. The title for the initial test was "The Beach on a Summer Day," and for the final test, "A Christmas Shopping Scene."

The Teaching Procedure.- During the training period of ten weeks, the pupils in the experimental schools were subjected to three different methods of teaching the writing of imaginative description; the method of picture study, the method of the literary model, the method of sense training. The pupils in the control schools received no direct training in the writing of imaginative description whatever.

Eliminating Variable Factors.- In order to be able to evaluate the one variable that was the chief subject of the study, i.e., teaching method, the investigator made every effort to eliminate all other factors that might have entered into and thus beclouded the conclusions. Pupil differences were reduced by having the various teaching groups equated on the basis of pupil intelligence. Teacher difference was minimized by having each lesson in each teaching procedure taught according to a definite and detailed procedure prepared by the experimenter personally, after conference with experienced teachers of English composition; and by rotating the teachers in such a fashion that no class was under the influence of a particular teacher more than a third of the training time. The fatigue factor was kept under control because all the teaching took place for all the pupils on the same day of the week and at the same hour of the day.

Method of Scoring the Tests.- The scoring of both initial and final tests was done by the experimenter himself on the basis of the scale values for imaginative content in Description set down in the Van Wagenen English Composition Scale. This scale was selected by the investigator after conference with his sponsor, a university professor of English and several teachers of junior-high-school and senior-high-school English composition, as being the best objective instrument for his purpose. Before undertaking the scoring task, the experimenter prepared himself carefully to the end that he might eliminate inaccuracy and any tendency to subjective basis.

- - - - -

11. Brooks, F. D., The Psychology of Adolescence (Boston: Houghton, Mifflin Company, 1929), 229.

Treatment of the Obtained Data.- Because the pupils in the Alpha and Beta Environments differed in intelligence, extent of equalization within the grades, socio-economic status, and racial-national background, the investigator concluded that it was best to consider the data as having been obtained from two parallel experiments in different environments. This gave him the opportunity for experimental as well as statistical confirmation or contradiction of the results obtained.

The data were first analyzed in an attempt to obtain light on the fundamental issue involved in the study; namely, which of the three teaching methods was most effective in improving the imaginative quality of pupils' descriptive writing? In addition to this main issue, the investigator sought information on such supplementary and related problems as the correlation between change and initial test, between change and mental age, between change and chronological age, between initial test and mental age, and between initial test and chronological age. An attempt was also made to discover differences in the accomplishment of the pupils that might be due to sex; and, by means of case studies involving the highest and lowest ranking pupils, to reveal what, if any, nonscholastic factors might have a bearing on the imaginative nature of pupils' descriptive composition.

V. Findings

The data secured by the experimenter were presented in tabular, graphic, and descriptive form. Briefly they are as follows:

With Regard to Effectiveness of Teaching Method

For the entire group in the Alpha Environment the order of merit was sense-training method, literary-model method, picture-study method. By grades in the Alpha Environment, the sense-training method showed distinctly better results in all grades, being followed (except in the 8B grade) by the literary-model method and the picture-study method, in that order. Similarly, for the entire group in the Beta Environment, the order of excellence was sense-training method, literary-model method, and picture-study method. No attempt was made in the Beta Environment to interpret the results by grades, because the presence of two variables (intelligence and teaching method) made definite conclusions impossible. In the combined environments, the sense-training method again secured better results than the literary-model and picture-study methods, which ranked second and third, respectively. In all cases and with all methods, the experimental groups, which had been subjected to direct teaching procedures, achieved an appreciably greater improvement than the control groups, which had had no direct teaching whatever.

With Regard to Correlations

(1) There is a negative correlation (-.68) between change and initial test, indicating that the higher the initial score, the smaller the amount of gain. (2) The correlation between change and mental age is positive but low (+.28), implying a slight tendency on the part of the more intelligent to gain more. (3) A positive but very low (sometimes near zero) correlation between change and chronological age appears to suggest little in common between a pupil's age and the amount of his improvement. (4) Likewise, a small negative correlation between initial test and chronological age implies little relationship between a pupil's age and his initial composition ability. (5) Between initial test and mental age, however, there is a fairly high positive correlation (+.46), leading to the presumption that there is a moderately close correspondence between a pupil's intelligence and his composition ability.

With Regard to Case Studies

The only conclusion of importance here was that the lowest ranking pupils appeared to have done little outside reading, while the highest ranking pupils appeared to have done considerable reading of this kind.

The correspondence between the extent of outside reading and the quality of imaginative writing was so persistent as to seem to warrant the inference that stimulating pupils to extensive reading should be an important feature of any program that is designed to improve the imaginative quality of pupils' written work.

With Regard to Sex Differences

The girls attained a higher average level of achievement in the initial as well as in the final test. But in the matter of amount of gain the data were not positive enough to warrant any definite conclusions.

VI. Educational Implications

For the Future Organization of Composition Teaching

(1) Composition must be studied intensively in order to isolate and evaluate the "intrinsic" qualities that enter into the composition product. Composition is too complex a process to be properly evaluated on the basis of general merit alone. (2) There is great need to set up and emphasize new values in composition teaching. Composition should be looked upon as the communication of our ideas to others; and these ideas, derived from the experience of our daily lives, are to a large extent based on our sense impressions. (3) Imagery should be an important element in composition. The present study seems to warrant the conclusion that a method of description teaching based primarily on sense training produces a better quality of imaginative writing than teaching based on more traditional methodology. (4) The school syllabus in composition teaching should provide for more training of an observational rather than of a linguistic kind. Since imagery appears to play an important part in the composition process, everything possible must be done to permit pupils to grow in power of observation in its widest sense. (5) Both teachers and pupils must be made aware that good descriptive writing transmits images by stimulating the imagination through the avenue of the senses.[12] No amount of bare detail is adequate for effective description. (6) Imagination requires training; it is not enough that it be left to itself; it needs definite stimulation and direction. (7) An intelligent program of composition teaching must be directed towards breaking down the barriers of restraint which hinder free and spontaneous self-expression. (8) The composition syllabus should make adequate provision for the acquisition by pupils of an effective vocabulary. The case studies in this research appeared to indicate that pupils who did poor descriptive writing were deficient in vocabulary ability.

For Use in the Classroom

(1) It is the duty of the teacher to create for pupils situations and help them to experiences from which imaginative expression may flow. (2) It is essential for the teacher to train pupils in the constant and intelligent use of their senses so that through them they may get full and accurate percepts. (3) The superiority of the sense-training method over the other methods employed in this research would appear to warrant the conclusion that first-hand experiences are more effective than second-hand experiences in developing imagination in composition writing. The teacher of composition should, therefore, stimulate pupils to direct and original sensuous acquaintance with the world in which

12. Tillett, N. S., and Yarborough, M. C., Image and Incident (New York: F. S. Crofts and Co., 1915), p. xi.

they live. (4) The teacher must work out a comprehensive program of vocabulary enrichment with his pupils so that they may have at command the necessary words to express their imaginative and emotional reactions to life's experiences. (5) The correspondence, as shown by the case studies, between reading and composition achievement seems to imply that the teacher must stimulate his pupils to extensive reading of good literature.

For Further Research

Questions implied in the present investigation, which future research might help to answer, are the following:

1. Was the generally ineffective work in the writing of imaginative description due to lack of descriptive vocabulary or to the absence of imaginative and emotional sensitivity?

2. Are the images used by pupils in their compositions real images, the result of their first-hand experiences, or are they merely verbal images, which they have acquired vicariously?

3. What effect does a pupil's nonscholastic background have upon his powers of written expression?

4. What is the relation between a pupil's intelligence and his composition ability from the point of view of content rather than form?

5. What changes in the imaginative quality of pupils' descriptive writing would result from a teaching procedure combining the various methods used in this study?

A STUDY OF THE FACTORS INVOLVED IN THE RATING OF PUPILS' COMPOSITIONS

by
Katharine L. Healy
The Johns Hopkins University

I. Purpose of the Study

This study was made in order to determine what elements influence teachers in their rating of pupils' compositions.

II. The Problem

Briefly stated, the problem of the investigation was to discover what seem to be the essential factors in compositions, evaluated by the teachers as excellent, superior, average, poor, and as failure.

III. Source of Material

The data on which this study was based included:

a. A collection of compositions secured through 50 elementary school principals from sixth grade teachers in their schools.
b. Two analytic charts which listed essential factors in compositions.
   (1) Analysis of quality.
   (2) Analysis of form.

IV. Methods of Procedure

As a first step, 175 sixth grade compositions were gathered from 50 different teachers. The five compositions from each teacher were to bear no comment other than the one word by which the teacher evaluated it.

a. One composition (of exceptional quality) valued Excellent.
b. One composition valued Superior.
c. One composition valued Average.
d. One composition valued Poor.
e. One composition valued Failure.

The second step was the rating of the 175 compositions by three trained judges who used the Nassau County Supplement to the Hillegas Scale (Trabue).

Third, a comparison of the composition was made for the purpose of indicating the range of grades. This means that all compositions were rated on the Nassau County Scale within their respective groups,--excellent, superior, average, poor, failure. This revealed differences in teachers' judgments. The problem was to discover where on the scale each Excellent composition fell, each Superior composition, and so on, through the several other evaluations. Also the study aimed to discover what range on the Nassau Scale was found among the five values.

TABLE I

COMPARISON OF QUALITATIVE RATING BY TEACHERS WITH NASSAU SUPPLEMENT SCORES ASSIGNED BY JUDGES

Points on Nassau Supplement Scale	Ratings According to Teachers' Judgments					
	Excellent	Superior	Average	Poor	Failure	Total
9.0						
8.0						
7.2	15	8				3
6.0	16	16	6			38
5.0	3	8	14	4		29
3.8	0	3	12	13	1	29
2.8	1	0	3	14	24	42
1.9				4	6	10
1.1					4	4
0						
Total	35	35	35	35	35	175
Median	6.8	6.4	5.2	3.8	3.1	

Fourth, an "Analytic Chart for Qualitative Analysis of Compositions" (Table II), prepared by Professor Florence E. Bamberger and the author of this study, was applied. The chart follows below.

The analysis chart preceding was for diagnostic purposes only, and was arranged for considering composition under each of five separate heads,--ideas, organization, self-expression, sentence sense, and vocabulary. With each head there are five graduated groupings of specific qualities which were regarded in reading each composition. Depending upon inherent qualities, a given composition received a grade of from one to five in quality.

A distribution of scores according to the Analytic Chart for Quality was arranged in order of frequency as the scores occurred on the compositions marked excellent, superior, average, poor, and failure respectively.

The problem here was to discover what elements found on the analytic chart were present in the Excellent compositions (of exceptional quality), in the Superior, and so on through the five levels.

Fifth, the compositions were rated to determine what proportion of those valued Excellent, contained "excellent" factors of the quality chart. Furthermore, what "excellent" factors were lacking? What factors other than those

TABLE II

ANALYSIS CHART FOR DIAGNOSTIC WORK IN COMPOSITION
Prepared by Florence E. Bamberger

Quality	5 Excellent	4 Superior	3 Average	2 Poor	1 Failure
Ideas	Original Deep Subtle	Interesting Clever New	True Clear Generally accepted	Few Commonplace Trite	Misleading Mistaken Untrue Lacking
Organization of Ideas	Effectively brought out point	Point well made Some suspense Relevant details	Point sensed but not skillfully made	Rambling, no plan but some sequence	Rambling No plan No sequence
Self-expression	Rapid Vivid details Brilliant	Fresh Vigorous Live	Reasonably coherent Commonplace Understandable	Vague Hackneyed Repetitions	Meaning obscure Unoriginal Labored
Sentence Sense	In harmony with the thought Vigorous Varied	Clear Smooth Mature	Pleasing Sequent Coherent Generally correct	Occasional violation of sentence sense Choppy Incoherent	Predominantly run-on type Predominantly phrase type Unintelligible Monotonous
Vocabulary	Fluent Vivid Wide Colorful	Well-chosen Varied Strong	Unconstrained Simple Accurate	Stilted Bombastic Lifeless	Offensive to good taste Narrow

TABLE III

ANALYSIS CHART FOR FORM

	Teachers' Marks on Composition				
	Excellent	Superior	Average	Poor	Failure
I. Language Correctness					
Verb form (errors)	3	7	18	29	45
Antecedent	7	4	12	6	10
Number	0	0	1	4	3
II. Correct Form					
Margin	0	2	2	4	3
Placement of Title	0	0	9	1	6
Indentation	2	1	3	2	9
III. Spelling (errors)	27	37	35	44	78
IV. Punctuation (errors)					
Period	0	0	3	1	10
Interrogation	1	0	2	0	5
Exclamation	0	1	1	3	0
Comma	38	21	39	36	33
Quotation	6	2	10	7	16
Possessive	5	3	7	4	10
Contraction	1	1	1	1	0
V. Capitalization (errors)					
First word of sentence	0	0	1	3	0
Pronoun I	-	-	-	-	-
Proper Noun	2	4	0	8	6
Proper Adjective	-	-	-	-	-
First word of quotation	0	1	7	9	7
Title of paper	3	6	7	11	9
Total Errors	95	90	158	173	250

considered as "excellent", were found, and in what proportion?

Sixth, a study of the compositions was made to determine what errors in form were found in the Excellent compositions and in what proportion. Similarly, form errors were determined in the compositions ranked as Superior, Average, Poor, and Failure (Table III). In the "Analysis Chart for Form" the factors have been arranged under three heads: (1) those termed "language correctness"; (2) those included under correct form; (3) those of spelling; (4) those included under punctuation; and (5) those included under capitalization.

Through the foregoing steps of studying the 175 compositions, a conclusion was reached as to factors which seem to have been essential ones in compositions evaluated by teachers as Excellent, Superior, Average, Poor, and as Failure.

Summary and Generalization

In conclusion, the study has brought out evidence to show: (1) The compositions tended to place themselves one step lower in quality evaluation, according to the diagnostic study, than they did in the teachers' general evaluation. (2) Teachers seem to distinguish among three degrees,--high, average, and low evaluation of compositions. They do not make a fine distinction when considering compositions under five heads. They do not make a distinction between an excellent (exceptional) composition and a superior one; nor do they make a distinction between average and poor compositions for the (sixth) grade. (3) The compositions marked failure by teachers, showed, to a marked degree, failure qualities in everything. (4) Finally, form seems to be a more potent factor to teachers and, therefore, receives from them more weight in evaluating compositions than does quality.

AN EXPERIMENTAL EVALUATION OF THE STYLISTIC APPROACH IN TEACHING WRITTEN COMPOSITION IN THE JUNIOR HIGH SCHOOL[1]

by
I. O. Ash
Shepherd State Teachers College
West Virginia

I. Purpose of the Investigation

The purpose of the investigation was: (a) to make a critical study of the stylistic[2] phases of composition writing and (b) to determine, as far as possible, the degree to which a knowledge of technical grammar contributes not only to the correctness of written expression, but also to the stylistic factors which insure a well-written composition.

II. Nature of the Experiment

The experiment was in the nature of a teaching project in the seventh, eighth, and ninth grades during the first semester of the school year 1929-30. Three high schools of the Eastern Panhandle of West Virginia participated in the study. One[3] was used as an experimental group and the other two as control groups.

Comparisons were made at the beginning of the semester to determine the similarity of the experimental and the two control schools. It was found that they were quite similar in that they were all district schools, organized on the six-year plan, and located in towns of about the same size. Their principals, teachers, and students were approximately equal in essential respects.

At the beginning of the semester the pupils were tested in both mental ability and writing skill. The Otis Self-Administering Test of Mental Ability was used for determining the mental ages, and for a composition test each student was asked to write for thirty minutes, using as the basis of his narrative a set of facts given him by the experimenter. The compositions were then graded by the use of standard composition scales. After the preliminary testing was finished, the teachers in the control schools went ahead with their work without any change whatever. They followed the state course of study and taught in their own way. Reports were secured from them from time to time in order to have some basis for comparison with the plans used in the experimental school.

In the experimental school the teachers of English laid aside their regular course of study and used a series of six units worked out by the experimenter. The six units included all the principles underlying effective written expression. In as much as the study was concerned primarily with the stylistic phases of composition and only secondarily with the mechanics of writing, the last two units, which dealt with grammar and the mechanics, were taught only incidentally. In reality, then, the work of the semester consisted of the study of the first four units, which are the only ones to be included in a later

1. This is an abstract of a doctorate thesis completed at the University of Cincinnati, August, 1930.
2. By stylistic factors is meant those that have to do with sentence structure, paragraph building, diction, unity, coherence, sequence, emphasis, clearness, exactness of expression, rhetorical fluency, beauty, creativeness, and the avoidance of a dead monotony of form.
3. The three schools are those at Shepherdstown, Harpers Ferry, and Berkeley Springs. The one at Shepherdstown was the experimental group.

section of this abstract.

In all three schools compositions were written once a week. To the teachers and pupils in the experimental school certain directions were given by the experimenter relative to the selection of topics, ways of correcting one's own theme, the motivation of writing, and the like. At the end of the semester the pupils in all three schools were given a second composition test in the same manner as at the beginning of the term. Then, in as much as the teachers in the experimental school went back at the end of the semester to the regular course of study, which was similar to those being pursued in the control schools, a third test in composition was given to the pupils of the experimental school three months after the close of the semester. The purpose of the test was to determine to what extent the principles taught were holding over.

III. Units of Study for the Experimental School[4]

A. First Unit: From a Stylistic Standpoint--The Sentence

High school pupils have much trouble to know just what a sentence is. The purpose of this unit is, therefore, to develop a working conception of it. Sentence unity is violated in five different ways:

1. Fragments of sentences wrongly used. Do not punctuate a part of a sentence--a phrase or a clause--as if this fragment were a whole sentence.

   Wrong: He lived in Salem. An old town by the sea.
   Right: He lived in Salem, an old town by the sea.

2. The run-on sentence. Be sure that too much is not included in one sentence, thus giving rise to a long, loosely connected, and rambling series of thoughts.

   Wrong: I found George at last yesterday, and he is a fine fellow, but not tne man for my work, and I regret this.
   Right: I found George at last yesterday. Although he is a fine fellow, I regret very much that he is not the man for my work.

3. The inclusion of irrelevant thoughts within the sentence.

   Wrong: George Washington, whose home was in Mount Vernon and who had red hair, rode every Sunday to Alexandria to church.
   Right: George Washington, whose home was in Mount Vernon, rode every Sunday to Alexandria to church.

#4. Unity in compound sentences violated by improper coordinations.[5]
#5. Difficulties in complex sentences.

B. Second Unit: From a Stylistic Standpoint--Among Sentences

In a composition all the sentences may be grammatically correct, but the total impression rendered ineffective because of similarity of sentence structure, poor transitions, and other violations of good usage. For effectiveness in any type of composition, therefore, one should strive for correctness relative to:

1. Sentence variety. Seek diversity by:
   a. Varying the kind of sentences as to purpose, structure, and arrangement of ideas.
   b. Varying the form of the sentence

4. These units are given in a very abbreviated form. In the thesis itself they are all worked out in detail.
5. All divisions marked by # were considered rather difficult for junior high school pupils and were not developed systematically in class.

as to length, position of phrases and clauses, the voice of the verb, the type of discourse, and the order of the subject, predicate, and object.

#c. Varying diction and expressional forms relative to appositives, the length of simple sentences, subordinate clauses, the use of verbals, and terminology.

#2. Coherence and sequence.

#3. Unity.

#4. Emphasis.

C. Third Unit: From a Stylistic Standpoint--Within the Sentence.

Watch carefully those factors dealing with:

1. Words.

a. Avoid the omission of small words, especially pronouns and prepositions.

Wrong: He has an uncle is very wealthy.

Right: He has an uncle who is very wealthy.

b. Do not repeat the same words unnecessarily.

Bad: Since several years had passed since the death of Lincoln, many people had forgotten him.

Better: Several years having passed since the death of Lincoln, many people had forgotten him.

#c. Avoid undue ellipses.

#d. Use simple, definite words instead of hazy, classical ones.

#e. Use words that are pleasing to the ear.

#f. Observe good usage.

2. The relation of words to each other.

a. Be sure that all pronouns have definite antecedents.

b. Be sure that all modifiers have something to modify.

c. Keep modifiers close to the words they modify and pronouns close to their antecedents.

d. Watch the word order closely to see that the exact meaning is brought out.

e. Strive for exactness of expression.

f. Strive for clearness.

g. Be as brief as possible.

h. Keep to one subject in each sentence.

i. Watch the agreement of the subject and predicate, guard against the wrong uses of collective nouns, challenge all adverbs, etc.

#j. Careful attention should be given to proper coordination and subordination.

#k. Strive for good beginnings and endings.

D. Fourth Unit: From a Stylistic Standpoint--Among Paragraphs (Compositions, Letters, Articles, and Books)

1. Keep in mind the various phases of good sentence structure.

2. Make an outline.

3. Strive for rhetorical fluency and excellence.

4. Try some creative writing.

IV. Results of the Experiment and Their Interpretation

A. Comparison of the Improvement Made by the Three Schools

The compositions written during the first semester by all the pupils of both the experimental and the control schools were scored on twenty-three different items. They are:

1. The number of words.
2. The number of "parts" punctuated as whole sentences.
3. The number of "stringy" or run-on sentences.
4. The number of sentences.
5. The number of declarative sentences.
6. The number of interrogative, imperative, and exclamatory sentences.
7. The number of simple sentences.
8. The number of compound sentences.
9. The number of complex sentences.
10. The number of compound-complex sentences.
11. The number of sentences beginning

with either a phrase or a clause.
12. The number of word omissions.
13. The number of undesirable repetitions of words, phrases, etc.
14. The number of antecedents that are not clear.
15. The number of modifiers that are not suitably placed.
16. The number of faulty agreements, such as those of verbs and their subjects, pronouns and their antecedents, adjectives and the nouns they modify, tense forms, etc.
17. The number of sentences the meanings of which are not clear.
18. The number of cases of faulty punctuation.
19. The number of cases of faulty capitalization.
20. The number of misspelled words.
21. The number of other grammatical mistakes, such as wrong compoundings, undesirable abbreviations, wrong forms for writing numbers, etc.
22. The mean form values of compositions according to the Willing Composition Scale.[6]
23. The mean content values of the compositions according to the Hudelson Typical Ability Composition Scale.[7]

In order to have some basis for grading on certain phases of sentence structure in compositions, such as the ratios of declarative to the number of interrogative, imperative, and exclamatory, the number that should be introduced by phrases or clauses, the relative number of simple, compound, complex, and compound-complex, the length, and the like, several paragraphs from well-known short stories by four modern and authoritative writers were read. It was found, for instance, that those writers began approximately twenty-five per cent of their sentences with either a phrase or a clause. Likewise other experimental evidence, such as that of Pressey's[8] relative to the frequency of errors in sentence structure, was taken into account in determining excellence in those matters wherein no absolute standards are available.

Because of lack of space, no attempt will be made, beyond that of a summarizing table, to give detailed results of any of the particular items on which the compositions were graded except the last two. On those two the following results were found:

Item 22 deals with the mean form values of the compositions according to the Willing Composition Scale. For grades seven, B[9] had a loss of 28 per cent. This is an interesting fact in as much as the teacher reported that she did but little else during the language period all semester except to drill in the fundamentals of grammar. School H had a small gain of 6.3 per cent. School S showed remarkable improvement by reducing its mean score errors from 7.1 to 4.69, or a total of 34 per cent.

In the eighth grade B again failed to make a gain or even to hold its record on the first set of papers. School H had a very slight improvement of less than .5 per cent. School S had the best mean score on both tests and made a very significant gain; it reduced its errors more than 24 per cent.

In grade nine both B and H had slight gains of 7 per cent and 2 per cent respectively. School S started with a much lower score than the other two and increased it almose 30 per cent. It seems

6. The form value of the Willing Scale is secured by dividing the number of grammatical mistakes in a composition by the number of words in the paper. Thus the smaller the quotient becomes the greater is the excellence of the theme.
7. The content scores are the values assigned the compositions according to the Hudelson Scale. The values of the scale range from a zero score, which is one of no value, to a score of nine. Note: The third set of papers were graded only on items 22 and 23.
8. Pressey, S. L., "A Statistical Study of Children's Errors in Sentence Structure," English Journal, XIV (September, 1925), 529-35.
9. The letter B will be used in the remaining part of this abstract to refer to Berkeley Springs School; H, to Harpers Ferry; and S, to Shepherdstown, which are the three schools that participated in the experiment.

TABLE I

SUMMARY OF PERCENTAGE GAINS MADE ON THE VARIOUS ITEMS DURING THE SEMESTER

Number of Item	Berkeley Springs			Harpers Ferry			Shepherdstown		
	Seventh Grade	Eighth Grade	Ninth Grade	Seventh Grade	Eighth Grade	Ninth Grade	Seventh Grade	Eighth Grade	Ninth Grade
2	69*	-9	44	11	50	70	89	72	78
3	-69	42	47	11	-8	-94	32	57	89
7	24	46	43	35	31	19	40	31	25
8	-139	-23	-39	-52	0	9	-11	-1	60
9	14	3	111	31	130	69	225	77	2
10	-12	-54	-30	-47	-40	118	250	27	-82
11	40	2	-40	-32	-33	14	-7	12	50
12	26	24	18	-3	48	-60	35	53	25
13	-400	0	67	23	0	0	39	90	77
14	25	50	69	15	42	45	55	71	80
15	0	0	50	0	0	0	66	0	100
16	20	46	25	19	14	-23	18	30	76
17	50	41	43	60	85	25	43	60	83
18	-8	-6	6	8	7	9	11	13	36
19	10	32	31	31	60	15	55	70	87
20	19	27	18	14	23	25	22	52	60
21	-25	-27	10	12	19	-8	18	-3	26
22	-28	-8	7	6	0	2	34	24	28
23	3	9	17	11	6	13	59	16	21

*These are all expressed in per cents. Those preceded by a minus sign are losses instead of gains.
NOTE:- Items 1, 4, 5, and 6 are not included. Item 1 is the total of words in all compositions for any one grade and is used only in determining the scores of some of the other items. Gains on Item 4 are such relative factors that it does not seem feasible to put them in quantitative terms. Since there was such a small amount of change in Items 5 and 6, they are omitted from the table.

fair to say that the gains made by S on the matter of form, or grammatical correctness, are very significant for the reason that it was not stressed at all in that school. In reality the teaching of it was not much more than an incidental matter.

Item 23 concerns the mean content values of the compositions according to the Hudelson Typical Ability Composition Scale. For the seventh grade we find that B and H both made slight gains of 3 per cent and 11 per cent respectively. School S began with a much lower initial score than the other two schools, but was able to reduce her errors almost 60 per cent. In the eighth grade B had a loss of about 8 per cent. School H gained 6.5 per cent, while S improved more than 16 per cent. The gains for the ninth grade are, in the order of B, H, and S, 16.9 per cent, 12.8 per cent, and 20.7 per cent.

To summarize[10] the results of all the twenty-three items on which the compositions were scored, with the exception of numbers 1, 4, 5, and 6, it seems perfectly obvious that S did much better than either B or H. In practically every case S had a better record than

10. See summary Table I.

either in mean scores, in improvement, or in both. Many of the gains made by S seem most significant and are far greater on the whole than those of the control schools. These conclusions hold not only for the gains made in the second set of papers over the results of the first test, but also for the improvement from the time of the second test to that of the third one.

There are three principal reasons for claiming significance for the experiment. In the first place, it must be kept in mind that one semester is a very short time to make any great change in composition ability. It takes years to develop good writers. In addition to the brief time during which the experiment was under way, it must not be forgotten that the scoring covered practically every phase of the principles underlying written expression. Had only an isolated item or two been taken for study, then the difference might have been far greater, but even then it is doubtful whether one could have expected more improvement than was made in some cases by S.

In the second place, as has already been stated, the gains that were made by S after the close of the experiment showed that there was considerable "hold-over" of the principles taught during the experimental period.[11] The pupils seemed to hold what they had gained and to add more to that improvement than the two control schools.

In the third place, the gains made by all three schools were treated statistically by means of the difference formula. The differences between the improvement made by the experimental school on the one hand and by the control schools on the other were significant. Although the results did not show absolute certainty in every case, still in nearly every comparison the percentages ran up well toward the 100 per cent mark.

B. The Relationship between Content and Form in Written Composition

Up to this point all comparisons of the results of the experiment have been between the progress of the pupils in the experimental school and of the pupils in the two control schools. As a secondary phase of this study, the data of the test papers of all three schools were thrown together for determining as far as possible to what extent a relationship exists between content and form in junior high school compositions. The scores on Items 22 and 23 were taken for this comparison. From a purely logical standpoint, it can readily be seen that the two factors of content and form cannot be entirely divorced. To illustrate their dependence upon each other, note the difference in the following sentence that a few marks of punctuation make:

The president of the board said the teacher is a fool.

"The president of the board," said the teacher, "is a fool."

Thus it can be seen that there is an interdependence of form and content, but it seems reasonably clear that that close relationship pertains most largely to mere correctness of expression. For such other factors of content as imagination, inventive skill, sense of humor, originality, logical power,

TABLE II

COMPARISON ACCORDING TO THE HUDELSON TYPICAL ABILITY COMPOSITION SCALE OF THE SECOND AND THIRD SETS OF PAPERS FOR CONTENT VALUES

School	Grade					
	Seventh		Eighth		Ninth	
	Second Composition	Third Composition	Second Composition	Third Composition	Second Composition	Third Composition
B	2.62	2.59	3.2	3.28	4.83	4.77
H	3.19	3.21	3.76	3.58	4.41	4.49
S	4.08	4.29	4.12	4.52	5.48	5.76

11. See summary Table II.

perceptive ability, feeling tone, etc., dependence of content upon form becomes far less than does mere correctness or exactness of meaning.

In this experiment 586 different papers were graded for both content and form. By means of the product-moment method a correlation between the two was worked out. It amounted to .59 ± .019. While this is not an exceptionally high correlation, it nevertheless is large enough to show that the two factors are dependent variables. Now, it is rather generally admitted that our schools emphasize form to the exclusion of sufficient content teaching. As a consequence the results have been very disappointing. Again, it should be remembered that, in this experiment, form elements were taught only incidentally in the experimental school but were stressed in the control schools. In as much as the experimental school made more progress, not only in content but also in form, it would seem that form must be considered the dependent variable. These facts, then, indicate that we have been placing the emphasis on the wrong variable. Instead of emphasizing form as we have been doing, we need to give more attention to the content of writing. So far as the results of this experiment are concerned, such a procedure would not only raise the content phase to a higher plane, but it would at the same time influence the form elements even more than if the grammar and its mechanics were stressed in the usual formal and non-applied manner. Since the elements of form are largely tools of correctness and exactness, it seems a logical conclusion that they become somewhat more dependent upon content than content upon them.

V. General Conclusions and Suggestions for Further Study

A. General Conclusions

The conclusions reached group themselves into two divisions. First, there are those proved directly and rather conclusively by the experiment itself, and, second, those derived by inference from the experimental and authoritative evidence of others, to some extent supported by evidence of this experiment, so far as it deals with such factors. The first group are termed "direct" conclusions and the second "indirect" conclusions.

1. Direct Conclusions

a. To apply the mastery formula of teaching to a unit type of curriculum means a great lessening in the amount of tiresome repetition of the same facts and principles year after year, and secures a higher grade of accomplishment than does the method of teaching without well-learned cycles of subject matter.

b. Many of the form factors of written composition can be secured by emphasizing content through its application in written thought. It pays to spend more time on the stylistic phases of composition and less on the grammatical. This is proved by the experiment itself. The experimental group, which was taught the stylistic and content phases with but little emphasis on form, made even more progress on form than did the control groups which had but little of the content phases.

c. The mechanics and formal elements have, aside from mere formal accuracy, but little relationship to effective expression. Not only is this true, but it has been shown in this experiment that stylistic phases of writing are largely independent factors, which, through their development, greatly affect the dependent elements of grammar, punctuation, and the like. This is proved to some extent by the comparatively high correlation between content and form together with the fact that, as previously stated, emphasis on content brought improvement in form. The many phases of formal grammar and punctuation improve in many respects without much definite attention to them as such. Clear thinking, which is a prerequisite for good expression, seeks an outlet through the medium of correct form. In short, much of the mechanics of writing is inseparably bound up with the content and is, in a way, a definite part of it.

d. The factor of individual difference is very forcibly shown in the data of this experiment. Even with those of practically the same educational opportunities, there were differences in expressional ability practically all the way from fluency down to an almost unintelligible type of thought.

e. In regard to many of the specific elements on which the compositions were graded, the data show that:

(1) Junior high school pupils cannot be depended upon to get rid of either the "part" or the "run-on" type of sentence. These errors still persist in spite of the efforts made to eradicate them.

(2) The length of sentences is comparatively short. As soon as the early adolescent begins using clauses, which add to the length of the sentence, his language errors begin to increase, largely because of confusion of thought.

(3) The junior high school pupil makes but little use of interrogative, imperative, and exclamatory sentences in his written expressions.

(4) But little use is made of clauses and phrases in the beginning of sentences. To make use of them is one of the best means of correcting the "and" habit and of securing a desirable form of sentence variety.

(5) The choice of words reveals a limited vocabulary. There are only a few pupils who show much concern for good diction. Not only is the diction poor, but fluent expressions are rare. Seldom does the teacher find a good, high-type, literary sentence.

(6) Simple and compound sentences make up a very large part of the sentence type, according to form.

(7) Considerable carelessness is noted in the matter of undesirable repetition of words and phrases.

(8) The number of antecedents that are not clear are few, primarily because of the infrequent use of complex sentences. Likewise the number of modifiers that are poorly placed are not numerous because of the great number of simple and compound sentences used.

(9) The lack of clearness of expression is not usually so great as to obscure the meaning entirely, but accuracy of thought is quite often lacking. Were a greater percentage of complex sentences used, the meaning would suffer far more than it usually does.

2. Indirect Conclusions

a. The amount of the formal elements needed as such by junior high school pupils is much less than heretofore supposed. Much time and energy have been wasted in trying to raise the level of all pupils to that of a technical knowledge needed by only a few writers.

b. The motivation of writing through the means of interesting, challenging situations will solve a great many of the perplexing problems of teaching composition. Since improvement is the result of effort expended with a conscious aim for that improvement, not of mere time spent or, in the case of composition writing, of the number of themes required, more satisfactory results are likely to be secured by a greater amount of writing of a highly motivated nature than by mere practice of a less stimulating sort.

c. More attention to individual needs, both of a diagnostic and a remedial nature, results in a greater reward by way of achievement and success. Individual needs can best be met through the laboratory method of teaching. Pupil direction in study is of fundamental importance, and the best form of it is secured under the conditions of teacher guidance in the problem-solving situations of a classroom or a laboratory. As compared with the recitation method of teaching English composition, the laboratory plan is much more worth while. Horner[12] found that the recitation method required nearly twice as much time as the laboratory plan and even then the former resulted in a somewhat inferior type of product.

12. Horner, Warren B., "The Economy of the Laboratory Method," English Journal, XVIII (March, 1929), 214-21.

B. Suggestions for Further Study

The conclusions of the preceding section imply needs for further study. Of the many that might be mentioned, a few that are crying loudly for solution are:

1. The laboratory technique of teaching written composition lacks the refinement of the procedure used in the science laboratory. Much needs to be done in an experimental way to determine the best forms of supervising writing in the junior high school.

2. There is no widespread agreement on minimal essentials in language. Many courses of study are still heavily loaded with dead and useless subject matter that should be replaced by curriculums based on use. Here is a fertile field waiting for cultivation by means of philosophic thought and experimental investigation.

3. The relative importance of essentials is still a mooted question. Is one grammatical error as detrimental as any other? What phases of punctuation, capitalization, and sentence structure should receive most attention and which ones less? Here, then, is room for additional study.

4. Courses of study need to be worked out by cycles and units. No investigations have yet been made on appropriate cycles of instruction in written expression. Much theory has been propounded on the unit type of teaching, but not much has been done in its application to composition writing.

5. There is much need of specific plans for cooperative teaching of English in the junior high school. The need exists not only for better plans but also for more persuasive eloquence in their behalf. High school principals and their teachers of content subjects are yet to be convinced, at least to the point of favorable action, that it is worth while to teach good English in all classes.

6. There should be further work on analyzing the specific characteristics and their "carriers"[13] in their different fields of written expression. We have long since learned that we cannot construct material things as wholes, but must make the separate parts and then fit them together. In composition writing, however, we have been trying the plan of securing humor without knowing what constitutes humor; of striving for beauty of expression without being familiar with its elements; or of including courtesy in our letters without knowing what constitutes it. A complete classification of specifics in the different forms of writing awaits the skillful experimental analyst.

13. The term "carriers" means the factors or elements that convey the various qualities of excellence in writing. For instance, one of the "carriers" of courtesy is the expression of good wishes.

EFFECTS OF IMMEDIATE AWARENESS OF SUCCESS AND FAILURE UPON OBJECTIVE EXAMINATION SCORES

by
C. L. Morgan and Lois V. Morgan
Purdue University

The main problem of this study was to evaluate the effect of immediate awareness of success and failure upon the results obtained from an objective examination. The general effect of immediate awareness of success and failure might be assumed: (1) to produce no appreciable modification in results; (2) to cause increased effort, attention, critical observation and thereby improvement; (3) to prove discouraging and therefore detrimental. A further problem was concerned with the retention of such learning as was made possible by the awareness of success and failure in comparison with gains made from test to test by the control group, which group had no objective criterion of success and failure. A third problem consisted of an evaluation of a proposed testing technique.[1] This technique consists of a self-scoring answer sheet which can be adapted to the various kinds of objective test items, e.g., true-false, multiple choice, matching, etc.; or it may be adjusted to provide self-scoring lesson assignments which could be incorporated in syllabus workbooks. For the present study, the test items and answer sheet were combined and printed. The test contained true-false columns in the left-hand margin except that a device now in the patent office, which cannot therefore be described in detail, allowed the student to determine immediately upon his answering of a question whether or not he had answered it correctly.

The experimental and control groups were selected at random from two classes in elementary psychology at Purdue University. Results from the Thurstone Psychological Examination were used to check the comparability of these groups. The experimental group contained forty-five and the control group forty-two students. In the last retest, the numbers of students within the two groups were reduced unavoidably to twenty and twenty-four respectively. The test items were composed of fifty true-false statements upon the problem of learning. The self-scoring test was given to the experimental group. The control group was given the same test in the usual mimeographed form. The initial test was given to both groups one week following the regular course examination which covered materials upon the topic of learning. The first retest followed the initial test three days later. The second retest came approximately two weeks later. The third retest was given fifteen weeks following the initial test.

In all of the retests, both groups were given mimeographed forms of the test. The items within each of the retest forms were presented in a new sequence. The initial test and the first retest were unannounced. The second retest was given during the final examination period. The third retest came during the second semester and was announced. In no instance, however, when the tests were announced, did the students anticipate being tested again over the same items. The influence of specific preparation must be assumed as a factor in the results of the second and third retests. Such influence in the initial test cannot be assumed, and it would appear to have been negligible in the first retest.

1. Sidwell, P., and Babcock, R. W. Self-testing sheet, patent applied for by Purdue University, 1933.

Both groups took the test at the same periods and were carefully proctored to prevent any kind of cheating. The scores made upon these tests were never reported back to either group. This was done as a control measure to prevent possible detrimental or beneficial effects in attitude of one group as against the other.

The results obtained from the Thurstone Psychological Examination, shown in Table I, indicated that the experimental group tended to be slightly superior in ability to the control group. This difference, however, was not completely statistically significant.[2]

TABLE I

GROUPS EQUATED FOR ABILITY UPON THE THURSTONE PSYCHOLOGICAL EXAMINATION

Groups	N	M ± S.E.$_{av.}$	S.D.
Experimental	42	51.50 ± 4.58	28.67
Control	43	45.98 ± 3.99	26.04
Difference	S.E.$_{diff.}$	D/S.E.$_{diff.}$	Chances in 100
5.62	6.07	.93	83

The results obtained from the objective examination, for the initial test and the three retests, for the experimental and control groups, are shown in Table II. The respective differences found for the two groups and their statistical significance are presented in Table III.

The results for the initial test indicated a slight tendency toward the superiority of the experimental group, as was found for the Thurstone Test, although again this difference is not completely statistically significant. The chances of a true difference on the Thurstone Test were 83 in 100, while for the objective test the chances of a true difference were 88 in 100. This would indicate that the slight superiority shown by the experimental group in the initial test covering learning should be attributed more probably to a slight superiority in ability rather than to the effect of immediate awareness of success and failure. Apart from this probable slight superiority in the ability of the experimental group, the two forms of the initial test (self-scoring and mimeographed) appeared to be equally effective as testing devices and to produce comparable results. The self-scoring test seems to have offered no advantages or disadvantages from the standpoint of initial testing.

TABLE II

AVERAGE SCORES OBTAINED ON THE INITIAL TEST AND ON RETESTS FOR THE EXPERIMENTAL AND THE CONTROL GROUPS

	1.Initial* Test	2.First Retest	3.Second Retest	4.Third Retest
Experimental Group				
Number	42	42	42	20
Mean	18.07	27.28	25.57	30.40
S.D.	7.47	7.96	10.02	11.70
S.E.$_{av.}$	1.15	1.23	1.55	2.62
Control Group				
Number	45	45	45	24
Mean	16.12	18.19	21.13	27.86
S.D.	7.49	8.34	8.01	8.31
S.E.$_{av.}$	1.12	1.24	1.20	1.74

*The initial test for the experimental group was self-scoring, while the initial test for the control group and all of the retests for both groups were mimeographed in the usual manner.

The retests, in spite of the limited samples, tended to indicate true differences in favor of the experimental group. The most striking increase found

2. Had the formula for correlated measures been used, the difference between these two groups would probably have been completely statistically significant.

TABLE III

SIGNIFICANCE OF DIFFERENCES BETWEEN THE EXPERIMENTAL AND CONTROL GROUPS ON THE INITIAL AND RESPECTIVE RETESTS

	1E-1C	2E-2C	3E-3C	4E-4C
Av. Diff.	1.95	9.09	4.44	6.54
S.E.diff.	1.61	1.75	1.96	3.15
D/S.E.diff.	1.21	5.19	2.27	2.08
Chances in 100	88	100	99	98

for the experimental group was upon the first retest. This appears to have been the direct result of the immediate awareness of success and failure in the initial self-scoring test. The experimental group tended to maintain their increased superiority, although, with the possible exception of the second retest for the experimental group, both groups tended to gain from one test to the next. The analysis of the gains from one test to the next, and the statistical significance of these gains are found in Table IV. The gains made by the experimental group, for the respective retests as compared with the initial test, were

TABLE IV

SIGNIFICANCE OF INTER-DIFFERENCES IN TEST SCORES FOR EXPERIMENTAL AND CONTROL GROUPS

	1*-2	1-3	1-4	2-3	2-4	3-4
Experimental Group						
Av. Diff.	9.21	7.50	12.33	1.71	3.12	4.83
S.E.diff.	1.69	1.93	2.86	1.98	2.90	3.05
D/S.E.diff.	5.45	3.89	4.31	.86**	1.08	1.58
Chances in 100	100	100	100	80	86	94
Control Group						
Av. Diff.	2.07	5.01	7.74	2.94	5.67	2.73
S.E.diff.	1.67	1.64	2.07	1.73	2.14	2.11
D/S.E.diff.	1.24	3.06	3.59	1.70	2.65	1.29
Chances in 100	89	100	100	96	100	90

*Initial test; other numbers indicate respective retests.

**Chances in a hundred of a true loss.

all statistically significant. The gains made by this group from the first retest to the following retests were not completely statistically significant, and tended not to be so great as were the respective gains shown for the control group. Upon the second retest, the chances were 86 in 100 of a true loss of the gains shown for the first retest. Persistent gains were made by the control group from the initial to the respective retests. These gains tended to become increasingly significant from the initial test to the succeeding retests. The gains made by this latter group, from any test situation to the next following, were not completely statistically significant. There is some indication that the two groups gradually tended toward their original status of equality with increased lapses of time.

To summarize the results, the self-scoring technique employed as a testing device appeared to produce no facilitating or detrimental effects. The retests indicated that the experimental group profited significantly through the experienced awareness of success and failure made possible by the initial self-scoring form of the test. This group also tended to maintain their increased superiority in later retests, although the two groups also tended to revert to their original status of equality with increased lapses of time. The maintenance of this gain probably resulted more from further study than from mere retention. The control group profited little from the initial test. However, they tended to gain consistently and relatively more rapidly from each test situation to the next, as compared with the experimental group. The persistent gain shown by the control group, rather than the maintenance of a quickly achieved increase in superiority, as was shown by the experimental group, was also more probably a result of continued study rather than improvement resulting from a repetition of the test, which test, when in mimeographed form, offered no objective criterion of success and failure.

It may be suggested on the basis of the results found in this study that

the self-scoring technique suggested by Sidwell and Babcock appears to prove equally effective as a measuring device to that of the mimeographed form of objective examinations, wherein awareness of success and failure is not apparent. In addition to presenting no difficulties as a testing device, this technique would serve as a rather effective teaching device. The method further reduces the difficulty of scoring in that no key is necessary. The accuracy of scoring would probably be significantly increased, while the scoring time would be radically reduced. If this method could be satisfactorily incorporated into a self-scoring syllabus work-book, it should prove effective as a teaching manual, and would reduce the amount and cost of paper work or clerical help.

In conclusion, the results indicate that immediate awareness of success and failure causes no significant change in the scores obtained from an objective examination. Immediate awareness of success and failure does produce a significant amount of learning as indicated by retests. The maintenance of such learning, without further study and review, probably represents a typical curve of forgetting; although in this study, where such review and further study did occur, the initial level of achievement tended to be maintained.

RIGHT: AN EXPERIMENT WITH AFTER-EFFECTS

by
Charlotte Harris
Belleville, New Jersey

After-effects of synaptic responses have long been recognized as a means of strengthening or weakening those responses. Corporal punishment, as a control of undesirable bonds, is almost synonymous with the little red schoolhouse. During recent years, however, much stress has been laid by mental hygienists upon the potency of fear in causing social maladjustment. Since fear is a consequence of an undesirable experience, we wonder if punishment, as such an experience, is necessary for learning. While corporal punishment has, of course, been prohibited in many states, sarcasm and other equally cruel forms of subtle punishment are still practiced.

There may be either pleasant or unpleasant after-effects of synaptic responses. If the pleasant could be proved strong enough so as to increase learning, educators might be prevailed upon to abandon the unpleasant, for the unpleasant after-effects, while preventing the formation of one bond, seem to cause a circle of harmful associations.

Thorndike, speaking of after-effects in his original law of learning, justified punishment by the following statement:[1]

> "When a modifiable connection between a situation and a response is made and is accompanied or followed by a satisfying state of affairs, that connection's strength is increased; when made and accompanied or followed by an annoying state of affairs, its strength is decreased."

However, in his recent book, Human Learning, he states:[2]

> "Annoyers do not act on learning in general by weakening whatever connections they follow..... Satisfiers seem to act more directly and generally and uniformly and subtly."

And in conclusion, he says:[3]

> "These experiments do not mean that punishment is futile. They need not predispose us to any change of attitude....But they will do so and probably should. Since in these experiments with these subjects, the wrong connections were simply displaced or nullified by the right ones, not intrinsically weakened, we may properly expect that something similar may happen in many sorts of learning, and we may increase our confidence in positive rather than negative learning and teaching."

The following experiment was performed for the purpose of observing speed of learning under two conditions: first, when the incorrect response is followed by annoyance; second, when the correct response is followed by satisfaction. The experiment was a vocabulary test set up in the following manner. From a list of five-hundred difficult words, fifty were selected because they had synonyms of one word, because knowing their meaning might be of value, and because they were probably

1. Thorndike, E. L. "The Laws of Learning," Educational Psychology, II. (New York: Teachers College, Columbia University, 1914), p. 4.
2. Thorndike, E. L. Human Learning, (New York: The Century Co., 1931), p. 46.
3. Ibid.

TABLE I

SAMPLE PAPER OF TEST I, ANNOYANCE, SHOWING METHOD OF SCORING

Trial 1
Mary Moe--8-II
13 wrong

Underline correct meaning:

1. Laudanum is a kind of: opium; insurance
2. Assiduity is: foolishness; diligence
3. Immolate means: vaccinate; sacrifice x
4. Temerity means: fear; rashness
5. A sapient man is: wise; devout x
6. Effervescent means: lively; gloomy x
7. Nebulous means: wicked; cloudlike x
8. Obsequious means: servile; dull
9. To disseminate is to: confuse; scatter x
10. Effigy refers to: image; speech x
11. Agog is: wicked; eager
12. Transcend means: surpass; cross
13. Impervious is: impenetrable; wicked x
14. Gordian is: friendship; intricate x
15. Inimical is: unfriendly; excellent x
16. Gregarious is: sociable; sad
17. Felicitate means: help; congratulate
18. Proletarian is: vulgar; single
19. Metamorphosis is: death; change
20. Inhibit means: hate; restrain x
21. An equestrian is a: hunt; horseman x
22. To matriculate is to: enter; insist x
23. A man's nemesis is his: fate; surname x
24. Lacerate means: weep; wound
25. Nonpareil is: unknown; unequalled

not part of the vocabularies of the subjects tested. These words were divided into two lists of twenty-five words each. The test was a dual choice one (see Tables I and II), the given word written first and two meanings suggested for it, one being the synonym, the other, a word selected either at random or because of a strong incorrect suggestion. There was no orderly arrangement; in the first test, however, thirteen wrong meanings were placed in the first position; in

TABLE II

SAMPLE PAPER OF TEST II, SATISFACTION, SHOWING METHOD OF SCORING

Trial 2
John Doe--8-I
17 right

Underline correct meaning:

1. Fervent is: earnest; alcoholic x
2. A mendicant is a: dressmaker; beggar
3. Zeal is: ardor; hope
4. Physiognomy refers to: face; body x
5. Icons are: images; verbs
6. Salient is: present; prominent x
7. Undulate means: wise; wobble x
8. To inundate is to: flood; invite
9. Vociferous is: volcanic; noisy x
10. An amanuensis is a: secretary; gem
11. Adamant is: firm; black x
12. To aver is to: curse; assert x
13. Astute means; cruel; keen x
14. Alacrity is: briskness; poverty
15. Volatile is: sharp; lively x
16. Affluent means: wealthy; loving x
17. Voluble means: fluent; poor
18. Sagacious means: arrow; wise x
19. Beatitude is: happiness; beauty
20. Banal is: sudden; wicked x
21. Scintillate means: vibrate; sparkle x
22. Obstreperous means: disorderly; inform x
23. Rigorous is: cold; stern x
24. Corroborate means: continue; verify x
25. To segregate is to: pray; separate x

- - - - -

the second test, thirteen correct meanings were placed in the first position.

The class chosen to take the test was an eighth grade of twenty-five students of average intelligence according to the results of The National Intelligence Test, Forms A and B.[4] Because of the absence of three people during the experiment, their scores have been omitted in the results, leaving a total of twenty-two persons taking the tests.

Tuesday, May 10, Test I, Annoyance, was given first, and then repeated four times until Monday, May 16, when retention was checked. Test II, Satisfaction, was given similarly, beginning Tuesday, May 17, and being repeated until Monday, May 23, when retention was again checked.

The procedure was as follows: the papers were passed face down, the class then being told to turn them over and listen to directions. For Test I, Annoyance, the directions were: "Write your name. Read each sentence carefully. Underline the synonym of the first word. If you are not sure, guess. Try not to make mistakes." About five minutes were allowed, but if a pupil was slow he was told to finish. When all had marked their papers, they were instructed thus: "I will read the word which is wrong. Check your answer if it is wrong." Then the instructor read, "Laudanum is insurance, wrong", etc. The papers were checked rapidly and immediately collected to prevent review.

The directions for proceeding with Test II, Satisfaction, were identical except for the last sentence which was, "Try to mark the right answer", instead of "Try not to make mistakes." The class then proceeded as in Test I. When all were finished the instructions for marking were, "I will read the word which is right. Check your answer if it is right." Then the instructor read, "Fervent is eager, right", etc.

The computations are based on the numerical average found by dividing the sum of the scores in a test by the number of students taking the test. In interpreting the results it was assumed that for each word guessed wrong one was guessed right. To find the average number of words known, the writer doubled the wrong responses and subtracted from 25. In each later trial, to find the corrected score, the investigator subtracted the number of wrong responses from the correct responses on that test; then subtracted the number of words known on the first test from this result. To find comparable grades for all tests, the ratio between the corrected score and the number of words not known on the first test was obtained.

When Test I, Annoyance, was given the first time, an average score of 14.1 was secured. This indicated that the class knew on an average 3.2 words. Hence the problem was to learn 21.8 words in order to make a grade of 100. Improvement was gradual. By the fourth test the grade was 60.6. On the retention test the raw score was 21.7, the grade 67.9, showing a slight improvement. The class was worried throughout the test in spite of the improvement and expressed dissatisfaction at the list. One child remarked, "When you say that a word is wrong, I remember it and get all mixed up."

The average score for Test II, Satisfaction, on the first trial was 14.3. According to our formula, 3.6 words were known, indicating that the first test was .4 of a word more difficult than the second; this is less than the probable error in measurement. The experimenter was surprised at this, for throughout Test II the class insisted that the list was much easier; their enthusiasm was so heightened that during the last trial some cheered when they found their scores. The learning in Test II was noticeably greater than in the previous test; the average grade in the fourth performance was 86.0. It improved to 86.9 in the retention trial. On the fourth trial, half the class reached 100 and maintained this grade on the retention trial. The

4. National Intelligence Test, Forms A and B, prepared under the auspices of the National Research Council. Yonkers-on-Hudson, New York: World Book Company.

TABLE III

DISTRIBUTION OF SCORES OF 22 STUDENTS OF FIRST EIGHTH GRADE ON 5 TRIALS OF TEST I, ANNOYANCE

	Trials				
	1	2	3	4	5*
Score					
1					
2					
3					
4					
5					
6					
7					
8	1				
9					
10	1				
11	3	1			
12	1	1			
13	4				
14	2	3	1		
15	2	4	1		
16	2	4	2	1	1
17	4	2	1		
18	1	1	4	2	
19	1	4	2	5	3
20		1	3	2	3
21			2	5	4
22		1	2	2	2
23				2	1
24			3	2	6
25			1	1	2
Raw average	14.1	16.3	19.6	20.7	21.7
Corrected score	0.0	6.4	11.0	13.2	15.2
Grade	0.0	20.2	50.5	60.6	67.9
Number of responses known - - - - - -	3.2				
Number of responses to be learned - - - -	21.8				

*Since the fifth trial was made on Monday of the school week, it tested retention.

members of the class were pleased with their results throughout and felt satisfied on the first performance, although an examination of the papers showed that there were more "less-than-thirteen-right" scores here than in the first test. In commenting on the test one child conceded, "While the first test is harder, the second is more tricky."

TABLE IV

DISTRIBUTION OF SCORES OF 22 STUDENTS OF FIRST EIGHTH GRADE ON 5 TRIALS OF TEST II, SATISFACTION

	Trials				
	1	2	3	4	5*
Score					
1					
2					
3					
4					
5					
6					
7					
8					
9					
10		1			
11	1				
12	3	1			
13	4				
14	6	2			
15	4	2	1		
16			1		
17	2	2	1		
18	1	2		1	1
19	1	1	1	1	1
20		2	3	1	1
21		3	3		
22			3	2	2
23		2	2	4	1
24		4	2	2	4
25			5	11	12
Raw average	14.3	18.8	21.6	23.5	23.6
Corrected score	0.0	9.0	14.6	18.4	18.6
Grade	0.0	42.1	68.2	86.0	86.9
Number of responses known - - - -	3.2				
Number of responses to be learned - -	21.8				

*Since the fifth trial was made on Monday of the school week, it tested retention.

- - - - -

This experiment was repeated in another group, but because of lack of time, each test was given only three times; otherwise the procedure was the same as before.

The group chosen for this experiment was another eighth grade, a class of twenty-four with high intelligence rating on the National Intelligence Test.[5] The group had been considered rather erratic and did not now belie its reputation. (See Tables V and VI.)

On Test I, Annoyance, the students knew an average of 2.6 words against 3.2 of the first eighth grade. On the third performance their grade, raw scores having been corrected, was 42.9, against 49.0. The low individuals were consistently low while the high were consistently high.

Test II, Satisfaction, was obviously easier, the initial raw average score of 16.3 indicating that 7.6, almost a third of the words, were familiar. The corrected grade on the final test was 49.4 showing an improvement of 6.5 over learning in Test I, which was not as significant as the former eighth grade's score of 68.2 on the third trial. Lack of interest, impotency of the after-effects, and copying may have blurred the results of the second eighth grade's attempt.

A more detailed idea of the experiment may be gained from a study of the tables. Only group data are presented, since the test was not long enough to furnish significant individual curves. There are a few general truths, however, that may be verified by the work which has been done.

Since the averages were for 550 units in each case (25 words given to 22 students) it is possible to draw conclusions for the tests as follows: Response followed by annoyance tends to create an unhappy class atmosphere. Response followed by satisfaction tends to encourage the group. There is learning whether the response is followed by annoyance or satisfaction. The learning is more rapid when response is followed by satisfaction. Therefore, both for economy in learning, and for social happiness, satisfaction as an after-effect is the better teaching device. Confidence in positive learning is well-grounded.

TABLE V

DISTRIBUTION OF SCORES OF 24 STUDENTS OF SECOND EIGHTH GRADE ON 3 TRIALS OF TEST I, ANNOYANCE

	Trials		
	1	2	3
Score			
1			
2			
3			
4			
5			
6			
7			
8			
9	1		
10	3		1
11		1	
12	2		
13	1	1	
14	9	3	3
15	2	4	1
16	4	2	2
17	1	2	1
18	1	4	2
19		4	4
20		1	4
21		1	1
22		1	2
23			1
24			
25			2
Raw average	13.8	16.8	18.6
Corrected score	0.0	6.0	9.6
Grade	0.0	26.8	42.9
Number of responses known - - - - - - -	2.6		
Number of responses to be learned - - - - - - -	22.4		

5. Ibid.

TABLE VI

DISTRIBUTION OF SCORES OF 24 STUDENTS OF SECOND EIGHTH GRADE ON 3 TRIALS OF TEST II, SATISFACTION

	Trials		
	1	2	3
Score			
1			
2			
3			
4			
5			
6			
7			
8			
9	1		
10	1		
11			
12	3	1	
13		1	1
14	1	1	1
15	1	1	1
16	5	1	
17	2	1	2
18		3	2
19	8	2	1
20	1	3	1
21	1	4	2
22		3	3
23		2	6
24		1	3
25			1
Raw average	16.3	19.1	20.6
Corrected score	0.0	5.6	8.6
Grade	0.0	32.2	49.4
Number of responses known - -		7.6	
Number of responses to be learned - -			17.4

THE DEVELOPMENT OF CLASSROOM PROCEDURES FOR TEACHING LARGE CLASSES IN SECONDARY SCHOOLS[1]

by
D. H. Stewart
Superintendent of Schools
Beaver, Pennsylvania

I. INTRODUCTION

A. Findings and Recommendations of Research on Class Size

Studies of class size have emphasized the need for a teaching procedure that is adapted to large classes because the use of the traditional recitational procedure and some of the recent general methods of instruction have proved to be inadequate. The national survey of secondary education revealed that traditional teaching methods persist in this country to a large extent. Statistics show that the increasing high school enrollment has been accompanied by an inincreasing class size. In view of economic conditions, it is unlikely that there will be a reversal to decreasing class size in the near future. Dr. Cooper,[2] former Commissioner of Education, summarized the results and meaning of studies of class size in the following statement: "It stands to reason that with large classes there must be some new techniques of instruction."

B. Purpose of Study

This study was conducted to determine the efficiency of instruction in a typical high school situation when all classes were large and when the subject plan of instruction was used to meet the problems arising in the teaching of large classes. The achievement of such classes was carefully analyzed and compared with the norms and records of small classes upon the same tests. The opinion of students and teachers was obtained in order to furnish evidence of the effect of such a program upon both.

C. Instruction

The subject plan of instruction[3] has been described in a number of studies which have been submitted to the University of Pittsburgh; in this study no attempt was made to refine the method. After studying this method the writer introduced it into the high school under his supervision in classes of the usual size. It was so successful that, when faced with other administrative problems, he recommended that it be given an experimental try-out with large classes. This recommendation was accepted by the school board and the program described in this report was put into operation.

Each pupil was given an overview of the subject in which the structural organization and the method of studying to be used were described. An attempt was made to have the pupil vision the entire subject; assignments were made with this in mind. Naturally it was necessary to make specific assignments upon blocks of the subject matter. These master guides were supplemented by core study guides for the different blocks

1. Summary of study presented to the University of Pittsburgh for the degree of Doctor of Philosophy. The study was directed by Dr. C. A. Buckner, Professor of Secondary Education, University of Pittsburgh. The program employed for the guidance of learning has been devised by Dr. Buckner.
2. Cooper, W. J. Economy in Education, Stanford University, California, 1933.
3. Theses and dissertations directed by C. A. Buckner at the University of Pittsburgh, and magazine articles written by him, describe the foundations of this plan.

and by oral instructions from the teacher. Core guidance was decreased from grade to grade during the year so that the pupil learned to be more self-reliant and less dependent upon the teacher. Not only was the pupil given a method of studying, but he was given an opportunity to test the efficacy of his method of thinking in the subject while he was in school.

Testing was integrated with learning and the gap between learning and testing, which usually exists, was reduced. The effect of testing upon learning became almost as meaningful to the pupils as the effect of learning upon testing. The form of the question was subordinate to the item to be tested; items were not segregated into true-false and other types of objective questions. Reports were made to pupils in terms of success with learning situations and not in terms of placement of scores in frequency distributions. By means of a self-appraisal technique for ability to answer and for accuracy of answer, training was given in control over knowledge.

The classroom was a place of real directed study instead of the usual type of recitation. Assigned oral recitations, reports, conferences, and directed study replaced the question and answer method. This type of class activity resulted in more definite, permanent, and meaningful learning.

The master study guides given the pupils were based upon learning threads. In history these were nine fundamental interests and activities of society; in English they were seven themes of literature upon which contributions may be made; in biology they were seven generalizations around which the course was organized to integrate zoology, botany, and physiology; in algebra conventions, principles, and laws formed these threads. Core guides based upon these were given the pupils as they began the study of different blocks of a subject.

This program of learning and teaching emphasized the assumption of responsibility by the pupil for his learning; the classroom became a workroom where the teacher guided and directed the learning activities of the students.

D. Description of the Experiment

1. Control and Experimental Data. Since all of the classes in the high school were large, it was impossible to have the usual type of controlled experiment. Fortunately, there were available in the files of Beaver High School scores on standard tests made in preceding years by small classes taught by many of the same teachers. These scores and the norms of the tests obtained from nation-wide testing were used as control data; the scores of the large classes constitute the experimental data. In addition the data reported by Symonds[4] in his study of ability standards for standardized high school achievement tests in American history, biology, and algebra 1 were used as control data. Wylie's[5] study in first semester algebra and Shank's[6] study in second year algebra were also used for control data.

In grade 11 of American history, English 4, biology, and algebra 2, Form A of the standard test was given in September, Form B in March, and Form A in May. In seventh grade American history, Form A was used in September and Form B in May. In algebra 1 a standard test was given at the end of each semester. In algebra 2 an additional standard test was given in May to supplement the main test used. In ancient history, in addition to giving two standard tests in May, three teacher-made objective tests were given

4. Symonds, P. M. Ability Standard for Standardized Achievement Tests in the High School. New York: Bureau of Publications, Teachers College, Columbia University, 1927.
5. Wylie, Carrie L. An Experiment in the Teaching of First Semester Algebra, University of Pittsburgh, Pittsburgh, Pennsylvania, 1932.
6. Shank, C. E. An Experiment in Directing Thinking in High School Algebra. University of Pittsburgh, Pittsburgh, Pennsylvania, 1930.

during the year. Scores for preceding small classes on these teacher-made tests were available. In problems of democracy a standard test was given in May.

2. Alternation of Subjects. In order to have large classes a program of alternation of the teaching of subjects was arranged. The median class size of Beaver High School was 44 which was much larger than the median for the state. Thirty was taken as the dividing line between large and small classes since that is the number recommended by the Pennsylvania Department of Public Instruction and the Middle States Association of Colleges and Secondary Schools.

3. Method of Obtaining Student Opinion. Two questionnaires were used. The first, which was given in May, was general in nature in order to avoid the suggestion of answers. Questions related to advantages and disadvantages of the program, subject suitability for large classes, and the desirability of the method of instruction. The second questionnaire which was given in November after the summer vacation and two months of study in the next grade, was a check upon the first. Through this questionnaire, evidence upon the consistency of pupil reaction was obtained.

4. Method of Obtaining Teacher Opinion. Ten teachers answered very fully sixteen questions about this program. These questions concerned the advantages and disadvantages of this program for teachers and pupils, the effect of this program upon class management and control, the value of clerical help and instructional supplies and textbooks, the applicability of the subject plan to large classes, the effect upon the health, disposition, and the attitude of the teacher toward teaching, and finally a frank estimate of the subject plan for either large or small classes. These answers of ten teachers to the problem of teaching large classes by the subject plan are very helpful in understanding the educational value of this program.

5. Adjustments to Meet Problems of Large Class Instruction. Clerks were employed to relieve the teachers of many of their clerical functions and to assist them in the preparation of instructional materials. The rooms were large enough to accommodate the number of pupils without undue crowding. There was no reduction in the budget for books or for instructional supplies. Every effort was made to give the program a fair opportunity and not to place any handicaps upon it because of lack of clerical assistance, textbooks, supplies, or overcrowding of the classrooms.

Table I presents a summary of the arrangement of the control and experimental classes and shows the record of their achievement upon the standard tests.

E. Mental Ability of Beaver High School Students

Data from the administration of three group tests of mental ability in the Beaver schools were available so that the control and experimental groups could be equated upon the basis of the intelligence quotient. By a careful analysis of the literature on the intelligence of high school pupils, and particularly that appearing in Monograph Number 4[7] of the recent survey of secondary education, entitled The Secondary School Population, it was established that the pupils of Beaver High School were typical of the usual high school population in ability. Some cause other than exceptional mental ability accounts for the achievement of the pupils in these large classes.

II. FINDINGS

A. Comparison with National Norms and Records of Small Classes

In every case the comparison between the achievement of the large classes and the norm of the standard test used

7. Kefauver, G. N., Noll, V. H., and Drake, C. E. The Secondary-School Population, p. 18. Bulletin No. 17, 1932, Monograph No. 4. Washington: Office of Education.

TABLE I

ACHIEVEMENT OF BEAVER LARGE CLASSES AND SMALL CLASSES

Year	Month	Class Size	Med. I.Q.	National Norm.	Q.D.	Median	Beaver Median in Q.D.	Median in Percentile
COLUMBIA RESEARCH BUREAU AMERICAN HISTORY TEST								
1934	May	41	106	81	23	96	0.65	66
1928	May	26	105	81	23	89	0.34	59
1934	Mar.	41	106	--	--	105	1.04	74
1931	Mar.	28	100	--	--	111	1.30	78
COLUMBIA RESEARCH BUREAU ENGLISH TEST								
1934	May	56	114	133	22	170	0.73	85
1928	May	31	109	133	22	149	1.17	68
1934	Mar.	55	114	---	--	153	0.91	73
1931	Mar.	23	106	---	--	159	1.18	78
RUCH-COSSMANN BIOLOGY TEST								
1934	May	55	107	40	11.5	59	1.65	84
1934	Mar.	55	107	--	----	44	0.35	60
1935	"	35	109	--	----	42	0.17	55
1931	"	30	97	--	----	36	-0.35	40
1928		30	105	--	----	36	-0.35	40
1926		18	107	--	----	44	0.35	60
COLUMBIA RESEARCH BUREAU ALGEBRA TEST SECOND YEAR								
1934	May	40	111	29	4	39	2.50	89
1933		18	111	29	4	34	1.25	78
1932		27	109	29	4	36	1.75	84
1934	Mar.	40	111	--	-	36	1.75	84
1933	"	18	111	--	-	28	-0.25	45
1932		27	109	--	-	36	1.75	84
1931		24	112	--	-	35	1.50	78
COLUMBIA RESEARCH BUREAU ALGEBRA TESTS 1 AND 2 FIRST YEAR ALGEBRA								
					Test 1			
1934	Feb. Pt. 1	60	113	19.8	7.0	27.2	1.06	83
	Pt. 2	60		13.3	4.5	15.3	0.44	67
	Total	60		33.0	9.5	42.3	0.98	81
1935	Feb. Pt. 1	47	119	19.8	7.0	26.0	0.90	83
	Pt. 2	47		13.3	4.5	13.7	0.09	55
	Total	47		33.0	9.5	40.0	0.73	75
					Test 2			
1934	May Pt. 1	60	113	11.0	3.5	15.90	1.40	89
	Pt. 2	60		6.0	5.0	18.75	2.55	91
	Total	60		16.0	7.0	34.50	2.64	90

favored the large class. This is shown in Table I. There are eleven comparisons between the achievement of the large classes in 1933-34 and the small classes in preceding years in Beaver High School. Nine of these, or 81.8 per cent, favored the large classes of 1933-1934. Since the subject plan was introduced in American history in 1930-1931, the fact that the small class record of 1931 in March excelled that of the large class in March 1934 might indicate that the subject plan gave better results with small classes.

B. Comparison with Symonds' Data

In eleventh grade American history, biology, and algebra 1, it was possible to compare the achievement of the Beaver large classes with those of the groups upon whose records Symonds[8] determined his ability standards for standardized high school achievement tests. The median scores of the Beaver large classes in these subjects excel those of the groups reported by Symonds. When the scores of the Beaver pupils are compared with the standard scores which Symonds set for pupils of different levels of ability, it is significant to note that the Beaver pupils at the lower levels of ability did proportionately better than those in the upper quarter of ability.

C. Achievement of Pupils with I.Q.'s Below 100

A careful analysis revealed that the mean score of Beaver large-class pupils whose intelligence quotients were below 100 excelled the norm of the standard tests with the exception of algebra 1 where an intelligence quotient of 105 was represented. The performance of pupils with intelligence quotients below 100 confirms the statement in the previous paragraph that pupils in the lower levels of ability are not handicapped by being in large classes when taught by the subject plan.

D. Some Significant Large Class Records

The most significant fact about the achievement of these large classes in biology is that in two heterogeneous sections, each drawn from grades 10-11-12 without regard to school success, previous work in science, or ability in or aptitude for science, the record of these large classes was better than any previous record in Beaver High School and better than the published norm. Such a performance under these conditions is especially noteworthy.

The large class in English not only excelled the median for public high schools but also the median score on the same test of the 1130 pupils who were graduated from 60 large Pennsylvania high schools in May, 1928, with marks of "A" in English. This class almost equalled the median score of students in private schools. The performance of large-class pupils was best in the vocabulary part of the test.

The achievement of seventh grade pupils in large classes in American history was exceptionally good. Their scores were obtained from two tests for two different years. In 1934-1935 the large eighth grade class that continued from the seventh grade had a median score equivalent to that of grade 11.5 in the history test of the Metropolitan Achievement Test, Advanced Battery, Form A. In ancient history the large class excelled the median of the standard test and equalled or excelled the record of previous classes on teacher-made objective tests.

In algebra 1 at the close of the first semester the large classes were slightly above the national standard. By the end of the second semester the large class was slightly better than public school standards in mechanics and had a median score three times as great as public schools in problem solving. This class excelled the record of private schools in problem solving. In algebra 2 the record was also proportionately better in problem solving. This is especially significant because most studies in algebra show that pupils are weaker relatively in problem solving than in mechanics.

8. Symonds, P. M. op. cit.

E. Summary of Student Opinion

From the pupil questionnaire evidence is presented to show that students felt that the subject plan placed the responsibility for learning upon them, developed better study habits, and gave them more varied experience. The chief disadvantages listed by the students were lack of individual attention and of immediate understanding of objectives. In meeting these objections the school has tried to bring the pupils to a better understanding of the meaning of assistance and has guided the pupils in making a real effort to solve their difficulties before asking for help. Further, an effort has been made to enlist the pupils' cooperation in promoting the work of the class period. In this experiment confusion was listed as a disadvantage in history, English, science, and mathematics in the order given. This would indicate that in subjects requiring numerous reference materials mimeographed instructions for the use of supplementary materials should be stated quite explicitly. Over 90 per cent of the pupils felt that the subject plan was better than the teaching method formerly used in Beaver.

F. Summary of Teacher Opinion

The teachers made some very definite statements that are in disagreement with much that is written about the attitude of teachers toward large classes. All felt that large classes can be taught successfully. Not one stated definitely that this program had any serious effect upon the health and disposition of those teachers assigned to large classes. Teachers agreed with pupils in stating that the chief advantage was the placing of responsibility for his learning upon the pupil and that the chief disadvantage was lack of opportunity to give pupils needed individual assistance. The most significant statement made by all teachers was that for large or small classes they preferred the subject plan to any other procedure with which they were familiar and that as class size increased the superiority of the subject plan became more apparent. These ten teachers would use the subject plan for large or small classes; it is the only method with which they are acquainted that could be used with large classes. These teachers stated that large classes presented a more challenging situation and that in meeting this challenge the teachers planned their work better with resulting benefit to the pupils. Although this evidence is subjective, it is worthy of consideration in reaching a final conclusion about the question of class size.

III. CONCLUSIONS

A. Summary

Large classes in history, English, biology, and algebra can be taught efficiently by the subject plan according to the evidence obtained in this experiment. Pupils of the lower levels of ability are not handicapped in large classes. The reaction of both pupils and teachers was favorable and differs from that which is ordinarily given.

B. Cost of Instruction

When taught in large classes the cost of instruction per pupil in grades nine to twelve, inclusive, was $52.00 and for grades seven and eight was $23.50. These figures are much below the average for communities like Beaver, and they are all the more significant when it is known that the average annual salary paid to the teachers of Beaver was $300.00 higher than that paid in neighboring districts. Large classes made possible a more efficient use of pupil stations. The Beaver High School building has many large classrooms, and unless classes are large, there is much unused space. Therefore the cost of operation and maintenance of buildings was also lower per pupil. These large classes made possible lower per pupil cost of instruction, higher average teachers' salaries, and better utilization of the building without lowering the standard of work in the classes.

C. Suggested Problems

This program is being continued, although the experiment has been completed. This indicates that from a practical

standpoint the policy of teaching large classes by the subject plan has been successful. Graduates from the high school who are now in college are reporting that they are not finding themselves handicapped but were helped materially by the subject plan. However, there are some problems that need further study. These are: (1) the development of better student-teacher relationship for individual pupil assistance, (2) the modification of the plan to afford more opportunity for the superior student, (3) the use of some of the financial savings of this program for higher salaries and for better instructional supplies, and (4) the development of a better understanding of responsibility for classroom activities.

D. Value of the Study

This experiment presents helpful information for administrators and teachers faced with the problem of maintaining desirable standards under difficulties of increased enrollment and decreased financial support. Since previous researches have indicated that a procedure adapted to large classes is needed, the data presented in this experiment make the subject plan of learning and teaching one which should be considered very carefully by any school system that is attempting to maintain and improve its instructional program.

- - - - -

MODES OF TEXTUAL PRESENTATION IN EDUCATIONAL PSYCHOLOGY

by
Alfred C. Horsch and Robert A. Davis
University of Colorado

The relative efficiency of various modes of presentation has been for many years an important topic in educational psychology. The problems investigated have consisted of a determination of the effectiveness of auditory, visual and other types of presentation, and in recent years there have been several studies dealing with the efficiency of silent and sound motion pictures. In addition to these, some attention has been given to textual, graphic and tabular modes of presenting quantitative data.[1] However, there seem to be no studies dealing specifically with the problem of determining the effectiveness of textual methods in presenting either theoretical or scientific material. The present study deals with the problem of determining the textual methods used by educational psychologists in writing their textbooks. It was believed that educational psychologists would reflect psychological principles in their own methods of presentation. The problem may be made clearer by raising several questions as follows: Do authors of textbooks stress general statements or do they make general statements and follow them with specific studies or examples and confirm the thesis maintained by other general statements? Or do they confine themselves to citation of specific studies or statements which tend to confirm the thesis of the topic or question discussed? Is the inductive approach which leads the student from the particular to the general more effective than the deductive approach which leads from the general to the specific? Or does the author use a combination of the two major methods of approach? Does he consider it best to treat certain topics entirely by means of specific statements when controversial issues are considered, thus allowing the student to form his own conclusions, or does he confine his treatment entirely to general statements when positive evidence is available? Obviously it is possible for an author to use predominantly any one of these techniques or combinations of several of them. This study was undertaken for the purpose of determining the techniques of presenting textual materials in educational psychology, as well as the consistency with which they are used.

Five textbooks in educational psychology were used for the analysis. Although these books were selected on the basis of their scientific nature, they do not all meet this requirement in the same degree; some are more advanced than others. This condition makes it possible to note variations in techniques as the books become more advanced. The following texts were selected for study:

Jordan, A.M. Educational Psychology.
Starch, Daniel. Educational Psychology.
Pintner, Rudolph. Educational Psychology.
Sandiford, Peter. Educational Psychology.
Trow, William Clark. Educational Psychology.

The methods of presenting subject matter in educational psychology may be divided into six techniques according to the arrangement of the general and specific statements.

1. Washburne, J. N. "An Experimental Study of Various Graphic, Tabular and Textual Methods of Presenting Quantitative Material," Journal of Educational Psychology, XVIII (1927), 361-76, 465-76.

1. General statements followed by other general statements.
2. General statements followed by specific statements and general statements.
3. General statements followed by specific statements.
4. Specific statements followed by general statements.
5. Specific statements followed by other specific statements.
6. Specific statements followed by general statements and specific statements.

Since a sharp distinction between general and specific statements cannot be made by using the criteria outlined in this study, controversial passages and topics may be included with equal justification under either classification. Specific statements comprise objective data of all types, descriptions of experiments, case studies, and illustrations of general statements, both hypothetical and real. When the style is constant, classification of topics under the various combinations of general and specific statements is easily made. Some authors, however, write figuratively and include many similes and short illustrations of one or more sentences. Since topics entirely devoid of short illustrations are seldom found and since it would be impossible to segregate systematically all of them, it has been necessary to classify arbitrarily those topics general in tone but with an occasional simile or short illustration as general.

The historical approach is sometimes used to trace the development of concepts, theories, and movements. Such treatment is considered general unless the investigations by historical authors are given in detail to illustrate the material they employ in developing conclusions. Psychology has not been sufficiently objectified to be considered a science throughout. Facts and theories not justified on the basis of objective study cannot be considered to have scientific value. This type of material is usually included under the general-statement technique.

Expert opinion is not conclusive evidence of the psychological advantage of one technique over another but it may form the basis for experimentation. The books selected were analyzed to determine the frequencies with which the various techniques of presenting material were utilized. Chapters, topics and books were then compared to note trends.

The chapter is chosen as a unit for purposes of comparison. Neither the chapters in the same book nor the same numbered chapters in different books are entirely comparable. The length of a book is not a valid criterion for the number of its chapters. Starch has 22 chapters with 450 pages while Trow, with the largest number of pages, 460, has the smallest number of chapters which is 11. It follows that the chapters in Trow will have more pages and topics than those of Starch. The same frequency of occurrence of a given technique of presentation for a given number of pages in Trow and Starch will appear different for the two books when the chapter is used as a unit; that is, the chapters in Trow will show higher frequencies than those in Starch.

Different subject matter is treated in the same numbered chapters in these books. The variation between the same methods of presentation in the same numbered chapters in different books will be greater since the type of subject matter has some bearing upon the frequencies. For example, a chapter on transfer of training will have considerably more experimental material than a chapter on the neurological bases of learning, which will alter the frequencies of the various topical methods of presentation.

However, the selection of a given number of chapters as a unit would involve an even greater number of obstacles. Chapters are logical divisions of books and trends may be noted in spite of the obvious defects in the technique of such an analysis.

In the comparison of the techniques of presentation for various chapters no consistency was found. In one

chapter one technique may be dominant, in the next chapter another technique, while in the third still another may be used. The technique of presentation depends largely upon the nature and importance of the topics under discussion.

A comparison of similarly numbered chapters in different books reveals no trend except a slight preponderance of general techniques toward the beginning of the books. Since most historical material and established theories and facts are treated generally this tendency is to be expected.

The topics consisting of general statements are significantly more numerous than any other group. They appear as frequently as the combined groups of general-specific, specific-general, specific, and specific-general-specific statements. Table I reveals considerable consistency among authors in the use of this technique of presentation. Jordan has a frequency of 40.5 per cent, Starch 41.0 per cent, Trow 43.7 per cent, Sandiford 46.6 per cent, and Pintner 56.0 per cent. With the exception of Pintner who treats approximately 13 per cent more topics by the general-statement technique, the average frequency is 43.0 per cent with a range of only 6.1 per cent. Since Pintner's textbook is probably the most elementary of all those studied, it may be expected that more topics would be treated by the general statement technique. This consistency may seem remarkable, since a comparison between chapters for each book shows a wide variation in the use of the general-statement technique. But the probable reason for the fluctuation of frequencies among the chapters is that, although a large percentage of topics is common to all books, they are treated in varying order. It may be concluded, then, that these authors consider a general treatment of topics the most important technique of presentation and that they agree to a marked extent on the topics that should be treated in this manner.

In contrast to the consistency of use made of the general statement technique, Table I reveals a wide variation in the frequency of the general-specific-general technique. Trow and Jordan are far above average--24.9 per cent--with frequencies of 35.1 per cent and 32.5 per cent respectively. Sandiford and Pintner are 10 per cent down the scale with percentages of 22.4 and 21.2, while Starch is in a class alone with a frequency of only 13.5 per cent.

TABLE I

COMPARATIVE AND TOTAL ANALYSIS OF TOPICAL PRESENTATION OF TEXTBOOKS BY AUTHORS

Method	Jordan		Starch		Pintner		Sandiford		Trow		
	Topics	%	Topics	%	Topics	%	Topics	%	Topics	%	Av.%
G*	101	40.5	74	41.0	148	56.0	138	46.6	138	43.7	45.6
GSG	81	32.5	24	13.5	56	21.2	66	22.4	111	35.1	24.9
GS	49	19.7	36	20.0	48	18.2	34	11.5	48	15.2	16.9
SG	13	5.3	18	10.0	6	2.3	35	11.8	15	4.8	6.8
S	4	1.6	26	14.4	6	2.3	22	7.5	2	0.6	5.3
SGS	1	0.4	2	1.1	0	0.0	1	0.3	2	0.6	0.5

*G - General statements
GSG - General-specific-general statements
GS - General-specific statements
SG - Specific-general statements
S - Specific statements
SGS - Specific-general-specific statements

The general-specific techniques again show greater consistency. The average frequency is 16.9 per cent with Starch, Jordan and Pintner grouped closely near the top; Trow and Sandiford fall behind at fairly regular intervals with frequencies of 15.2 per cent and 11.5 per cent respectively.

The specific-general, specific, and specific-general-specific techniques rank considerably farther down the scale, with the frequencies of the specific-general-specific technique almost negligible. For the specific-general technique, Sandiford and Starch are grouped closely together with 11.8 and 10.0 per cent, while Jordan and Trow show only half this frequency with 5.3 per cent and 4.8 per cent, and Pintner follows with 2.3 per cent.

With these figures in mind the more advanced and objectively treated textbooks of Starch and Sandiford may be analyzed. The general technique may be disregarded for the present, because it is usually employed with subject matter not strictly scientific. The specific-general-specific technique need not be considered because of its low frequency. It is shown, then, that these books rank lowest in the general-specific-general and the general-specific techniques and highest in the specific-general techniques and highest in the specific-general and specific techniques. However, on the basis of frequencies there still is found a somewhat larger percentage under the first two groups. Starch has 33.5 per cent in the first two groups and 24.4 per cent in the second two groups, while Sandiford has 33.9 per cent and 19.3 per cent respectively. In brief, approximately 60 per cent of the topics in Starch and 64 per cent of those in Sandiford belong to the general-specific-general and general-specific groups. Since the general-specific-general technique implies a certain amount of inductive thinking it may be concluded that approximately one-half of the topics are presented inductively in the most objectively treated textbooks. When the other textbooks are examined, deductive treatment of a large majority of the topics is indicated. On the basis of these facts it may be concluded that as textbooks become more objective they become increasingly more inductive in method of presentation. In other words, the presentation follows more closely the reasoning process by which the authors themselves deduced conclusions. This finding would indicate that authors consider only advanced students capable of inductive reasoning. Is this true or would greater stress on the inductive method benefit the beginning student? Here a fertile field of investigation presents itself.

Starch and Sandiford make comparatively wide use of the specific statement technique, having frequencies of 14.4 per cent and 7.5 per cent respectively. Are they justified in their assumption that the student, without guidance, will form logical conclusions on the basis of stated evidence; or are these authors complicating their treatment with raw data which are meaningless to the student without some interpretation?

Table I shows clearly the popularity of the deductive approach. The general-statement techniques constitute almost one-half of the total number. The general-specific-general technique is most popular for the presentation of strictly scientific material but there is wide disagreement concerning the extent to which this technique should be used. The specific-general-specific technique cannot approximate the integrative thought process of the author. It leaves "loose ends" which probably prove confusing to the student. The fact that it is rarely used indicates the opinion in which it is held by the authors. It might be advantageous, however, to experiment with this technique for the purpose of determining the extent of the effect of modes of textual presentation upon the learning process.

Analysis of the techniques of chapter presentation indicates that the objectivity or degree of scientific development or the type of student for whom the text is written appears to have no relation to the techniques of chapter

presentation. Starch develops his chapters deductively, which is in marked contrast to his stress upon the inductive approach in his topical presentation. Trow is consistently deductive in both instances, although he stresses the general-specific-general techniques in topics and the general-specific technique in chapters. Pintner uses the general-specific-general techniques in every chapter. Since Jordan uses the specific-general technique almost entirely, his method is inductive. Sandiford rarely uses the problem or the summary; therefore, his chapter presentation is specific.

This wide variation among authors, together with their own variation in topical and chapter presentation, reveals the need for experimentation to test the relative efficiency of the techniques of chapter presentation. The fact that an author favors a specific technique for topical presentation is no indication that he will favor that technique for chapter presentation. It is possible that the difference in the length and number of divisions in a book makes a different approach desirable for these two units.

General Conclusions

The study reveals several trends in the writing of educational psychology textbooks. The deductive approach is favored by all of the writers under consideration. More than 85 per cent of the topics are treated by the general, general-specific-general and general-specific techniques. Caution must be observed, however, in the interpretation of these results, since a large part of the treatment is not strictly scientific in character. The subjective data of experience, the historical approach, and the increased effort of the elementary textbook writers to relate subject matter to the everyday experience of the beginning student are not amenable to scientific treatment. These types of subject matter have been arbitrarily classified as general, and when they are eliminated the general statement technique is not adaptable to the development of logical conclusions.

The general-specific-general technique was found to be the most popular method of presentation of strictly scientific material, but wide disagreement exists as to the extent to which it should be used. This difference in opinion indicates the need for experimental data which will show the relative effectiveness of this technique for elementary as well as for advanced students and for various types of material. The general-specific technique ranks second in importance and is used with greater consistency.

When the more scientifically treated textbooks are examined, it is found that greater emphasis is given to the inductive approach. The specific-general and the specific technique come into greater favor. The inductive method is used almost as frequently here as the deductive method. The question of importance to scientific writers is the extent to which they can safely go in this direction.

The specific-statement technique is used almost as frequently in the advanced texts as the specific-general techniques. Are authors justified in presenting material specifically with the assumption that the student will, without guidance, form logical conclusions, or are these authors complicating their treatment by inserting raw data which are meaningless without interpretation? The analysis shows that the advanced student is assumed capable of forming his own conclusions, while the beginning student is not.

There are other questions which can be answered only by experimentation. Assuming that the several techniques are of equal value, is a consistent use of one technique throughout the book as effective as a variation of several techniques? It may even be shown that a somewhat inferior form of presentation may be used advantageously if the need for variety proves to be an important factor.

Do the same factors which determine the merits of a topical technique of presentation prevail as criteria for the technique of chapter presentation? It has been shown that there is little correspondence among authors in treating these

two units. It is probable that the difference in the length and the number of divisions of a book makes a different approach desirable for these units of subject matter.

Another question is the relation between types of material and the various techniques. It is probable that in a field where investigations are numerous but contradictory the general-specific-general technique will be more effective for purposes of evaluation, while the general-specific or the specific-general statement may prove more effective for more conclusive evidence.

SOME FACTORS IN TEACHER LEARNING

by
M. R. Dodd, Assistant County Superintendent
Kanawha County Schools
Charleston, West Virginia

Trends of Previous Research

There are observable differences between good and poor teachers but the causes of these differences are not easily discernible. Of fourteen studies[1*] reviewed, eleven revealed an average correlation of .158 between intelligence and teaching success; eight found an average correlation of .287 between scholarship and teaching success; and three gave an average correlation of .529 between personality and teaching success, but there is slight evidence that any one phase of personality may be used to forecast teaching ability. Success in practice-teaching, while of some value as an indicator, does not always portend success in public school teaching.

In an analysis of the teaching and background of 262 individuals, Kriner[2] found a correlation of .65 between teaching success and length of prior interest in teaching.

While as yet no satisfactory means of forecasting success has been discovered, the many previous studies have indicated the difficultness of the problem and doubtless will result in improved technique in future studies. A. S. Barr[3] is at present directing a series of studies as to how teachers learn, which promises to be more fruitful than any thus far conducted.

Studies reported by Kyte,[4] Pittman,[5] Collings,[6] Hoppes,[7] Sherwood,[8] and Brueckner[9] indicate that teachers improve decidedly under various types of supervision.

Melby[10] found a great divergence between the practices actually employed by supervisors and the practices which teachers and experts consider of most value. Among those activities assigned highest value by both teachers and specialists are demonstration teaching, individual conferences, special studies by individual teachers, and teachers' meetings where specific problems are studied. His data reveal the fact that none of the thirteen supervisory activities rated as being of greatest value by the majority of both teachers and specialists occurs among the four most widely used activities.

Valentine,[11] Kyte,[12] Leman,[13] Pickett,[14] and Fuda[15] find that supervisory procedures which enable the teacher to come into direct contact with a specific problem under the guidance of the supervisor are most frequently assigned the highest value. All five assign primary importance to individual conferences for improving teaching.

Fuller[16] has reported a carefully planned and executed study in which the relative effectiveness of individual and group conferences for teacher learning was evaluated. The data sustantiate the value of supervision according to five tabulations. The superiority of the individual supervisory conference over the group conference for improving the teaching of reading "is at least three to one as shown by twenty measures of difference" and for arithmetic "is at least one and one-half to one as shown by twenty measures of difference."

The distribution and content of the practice teaching course in the Salisbury, Maryland, State Normal School was made the subject of a study reported by Edna M. Marshall,[17] the data of this very carefully planned study justify the

*Refers to the bibliography at the end of this report.

following conclusions:

"In general, results of the study tend to show:

1. That the long period of practice is desirable.
2. That a given number of hours of student-teaching graded and distributed over a period of practice is more effective than concentrated practice of the same number of hours.
3. That a given number of hours of student-teaching graded and distributed over a period of practice is more effective than the same number of hours divided between directed observation and actual teaching."

While teachers colleges exhibit no great uniformity in their facilities for teacher training, there are certain procedures which have been generally accepted as requisities of any effective school program.

Three recent studies, by Jarman,[18] Flowers,[19] and Voelker[20] using a checklist survey and expert evaluation, indicate that authorities are generally agreed as to the value of student-teaching as a necessary part of the training-school curriculum. All three agree that both observation and practice are essential. Jarman and Voelker feel that the program should provide three distinct phases--observation, participation, and practice-teaching in two subjects.

All agree that the number of student-teachers to each supervisor should be such as to insure ample opportunity for individual conferences. Flowers believes that not more than one student-teacher should be assigned a training teacher at one time. Voelker recommends a total supervisory load of approximately ten student-teachers. Jarman found a recommended average ratio of one supervisor to eight student-teachers. This is the same as recommended by Learned and Bagley.[21]

Flowers indicates that since experimental and practice schools have different purposes, they should be maintained separately. Jarman states: "The campus school should be supplemented by one or more off-campus schools closely affiliated with the university for the purpose of furnishing additional directed teaching facilities under more typical conditions."[22]

Before any supervisory procedure becomes a stable practice either in the teachers colleges or in in-service supervision it should be validated by controlled experiment. While evidence tends to justify certain procedures, these are few in number, and possibly none is supported by conclusive evidence.

The absence of any standard conception of the amount and kind of student-teaching, observation, and participation desirable in the preparation of teachers is indicated by the wide range in practice as reported by investigators. The American Association of Teachers Colleges[23] reported a range of 2 to 600 hours for the colleges of their membership in 1925. Armentrout[24] reported in 1927 a range of 54 to 306 for 42 state teachers colleges.

It is natural that most institutions will adopt the generally accepted practices while waiting for more facts in the case. However, the obligation of justification rests on those who employ the procedures. This viewpoint is well stated by Hudelson: "The tradition that human behavior is hard to predict, control, and measure does not render it wholly resistant to scientific treatment; nor is the fact that education involves subjective, imponderable elements an excuse for neglecting to reduce to a scientific basis every element that can be made to lend itself to such treatment."[25]

Rankin concludes the first chapter of the Seventh Yearbook of the Department of Supervisors and Directors of Instruction thus: "The conclusion is that good supervision entails the use of the scientific method to an ever increasing degree. The potentialities far outweigh the limitations. Progress in supervision, and in the general attitude toward supervision, will come as supervisors possess the scientific attitude and meet their problems by the application of the principles of scientific procedure."[26]

The value of practice in the

training of teachers is almost unanimously recognized. There is convincing evidence of the value of some type of supervision, but there is a dearth of objective data as to what factors in the supervisory program accelerate learning and what factors have a disparaging influence or are ineffective. With the advent of scientific method pragmatic imposition of empirical and adventitious procedures cannot long be justified.

New processes and techniques may be evolved, but the selection of the best from the present abundance is "a consummation devoutly to be wished." Believing that an analysis and evaluation of generally accepted supervisory procedures is likely to be the most immediately fruitful next step, the writer has undertaken the present study.

The Problem

During the year 1931-1932, the writer conducted a study at West Virginia University Demonstration High School to determine the influence on teacher improvement of some factors in the supervision of student teaching, particularly the influence of specific training in questioning.

To evaluate the influence of training in questioning, each of six supervisors had one experimental and one or more control classes. In addition to the regular supervision, the experimental groups of practice teachers received special training in questioning.

Teacher improvement was used to measure the efficacy of supervision. The criteria for evaluating teacher improvement were: (a) pupil achievement, (b) rating of the teacher by committees of specialists, (c) lesson plans, and (d) Boardman's Procedures Test.

The data of the study are based on the activities of six supervisors, twenty-eight practice teachers, and 370 pupils of thirteen classes in English and social studies.

Pupil achievement was measured by standard tests. From each supervisor's two classes, paired groups of pupils were chosen for the purpose of securing data on pupil achievement. These pupil-groups were equated on the bases of age, I.Q., average mark in high school subjects, and score on a standard test in the class subject. At the same time the entire enrollment of all classes in the experiment participated in all tests, and progress for classes was tabulated as well as for paired groups. The data relative to scholarship were secured from the administration of the following tests:

I.Q. - Otis, S. A., Tests of Mental Ability, Intermediate Examination, Form A.

English 10 -
Tressler, English Minimum Essentials Test.

English 7 -
New Stanford Language Usage Test.

Social Science 10 -
Hill, A Test in Civic Information.
Hill, A Test in Attitudes.

Social Science 11 -
U. S. History, Pressey-Richards, A Test for Understanding of American History.

Home Economics -
Engle-Stenquist, Home Economics Test.

The educational tests employed in the pairing of pupils at the beginning of the experiment were administered again at the close of the study. The gain in score served as a criterion of teaching. The initial tests were given the first week of February, the final tests the last week of May. Thus the experimental period was slightly less than four months.

A summary of the comparative data for the two groups of student-teachers is given in Tables I and II.

The following formulas[1] were used in these and other similar calculations throughout the study:

1. Garrett, Henry E. Statistics in Psychology and Education, pp. 27, 121, 126, 129. New York: Longmans, Green and Company, 1930.

TABLE I

SUMMARY OF PAIRING DATA FOR STUDENT-TEACHER GIVING DIFFERENCE IN MEAN FAVORABLE TO EXPERIMENTAL GROUP, S. D. OF THE DIFFERENCE, AND CRITICAL RATIO

	Experimental Mean	Control Mean	Diff.in Favor Exp. Group	S. D. Diff.	Critical Ratio
Age	21.5	23.7	-2.2	1.1	2
Honor Points	1.42	1.62	-.2	.14	1.4
Honor Points 1st Subject	1.44	1.65	-.21	.166	1.3
Hours 1st Subject	36.9	37.1	-.2	2.69	.07
Hours Education	15.86	17.7	1.84	2.47	.7
Procedures Test	30.8	31.1	-.3	1.35	.2

TABLE II

SUMMARY OF PAIRING DATA FOR STUDENT-TEACHER GIVING DIFFERENCE IN MEDIAN FAVORABLE TO EXPERIMENTAL GROUP, S. D. OF THE DIFFERENCE, AND CRITICAL RATIO

	Experimental Median	Control Median	Diff.	S. D. Diff.	Critical Ratio
Age	21	22	-1.	1.37	.73
Honor Points	1.48	1.48	0	.176	0
Honor Points 1st Subject	1.55	1.50	.05	1.98	.25
Hours 1st Subject	35.5	35	.5	3.3	.15
Hours Education	15	15	0	3.1	0
Procedures Test	32	32	0	1.66	0

$$\text{S. D.}_{(\text{dis.})} = \sqrt{\frac{\Sigma FD^2}{N}}$$

$$\text{S. D.}_{(\text{av.})} = \frac{\text{S. D. (dis.)}}{\sqrt{N}}$$

$$\text{S. D.}_{(\text{mdn.})} = \frac{5}{4} \cdot \frac{\text{S. D. (dis.)}}{\sqrt{N}}$$

The reliability of differences between groups was calculated by the following formula, substituting "median" for "mean" or "av." where necessary:

$$\text{S. D.}_{(\text{diff.})} = \sqrt{\text{S. D.}^2_{(\text{av.1})} \qquad \text{S. D.}^2_{(\text{av.2})}}$$

$$\text{Critical ratio} = \frac{\text{Diff. between mean}}{\text{S. D. of the difference}}$$

The difference between means for the various factors on which student-teachers were equated is in no case very significant. According to Garrett[2] a ratio of 3 between a difference and the standard deviation of that difference is necessary to insure an actual difference.

All differences that do exist are in favor of the control groups. The critical ratio of 2 for age is the most

2. Ibid, p. 133.

nearly significant and indicates that the chances of an actual difference are 98 in 100.

In a small group such as this the median is probably a better measure of central tendency than the mean. It will be noted in Table II that the medians are identical for three of the pairing factors and exhibit a very small difference in the other three factors.

Using the median instead of the mean, the ratios of difference to standard deviation of the difference become quite insignificant. With three factors the critical ratio is zero. In the factors of "total hours teaching subject", "honor points earned in teaching subject courses", and "age", the ratios are .15, .25, and .73 respectively. A ratio of zero means that the chances of a true difference as small as zero are 1 to 1. For a ratio of .15, the chances are a little more than 1 to 3.

There were twelve groups of high school pupils participating in the study. Six of these, one with each supervisor, were taught by practice-teachers receiving special instruction in questioning and were thus designated, "experimental." The other six served as "controls."

There was no significant difference at any point in the initial data for the paired groups of pupils. While the experimental group had a slight composite advantage in age, the control group had the advantage in I. Q., average grade, and score on the initial test in subject matter.

Since the high school classes participating in this study were distributed among several school departments, it was impossible to measure pupil-progress with identical tests. Because with each change of subject a different test was necessary, five different tests were employed. Each of these had different norms for the grade in which they were used, and a semester's progress was in no two tests the same per cent of the initial score.

In order to have a more scientific basis for comparing progress, a weighted score was obtained, by dividing the actual gain by the standard deviation of the distribution of the scores on the initial test. Table III gives the initial and final scores on the educational tests, together with the actual and weighted gains for both the experimental and control groups.

It will be noted that the gain or loss was not consistent. In Group F, for example, there was a gain by both experimental and control classes, while in Group A there was a loss in all sections. In Group C, the experimental section of paired classes made the most progress, while the control section of the unpaired classes made the greatest gain.

However, the average of all sections, shows a greater gain for the experimental group. The paired pupils show a weighted gain in score of 3.1 in the experimental sections and a gain of .1 in the control, (Table III).

While neither the actual gain nor the weighted gain is statistically significant in any one of the classes, it is portentous that the gains are uniformly favorable to the groups receiving training in the use of the question.

Measuring Teaching Success by the Ratings of Observation Committees

As a further measure of improvement each practice-teacher was rated by a committee of specialists at the beginning and again at the close of his practice-teaching period. Both the initial and the final ratings of each teacher were made by the same committee. The rating committees were composed of individuals selected on the bases of maturity and experience from two graduate classes in supervision at West Virginia University. The members of these committees spent several hours with the writer in discussion and use of the rating card before attempting the official ratings.

A summary of the evaluation of teaching progress as measured by these judges is given in Tables IV and V. The ratings indicate, with few exceptions, a general gain in "personal qualities",

TABLE III

PUPIL PROGRESS IN VARIOUS SUPERVISORY GROUPS OF PAIRED PUPILS SHOWING PER CENT GAIN DURING THE EXPERIMENT

	1st. Test Av.	2nd. Test Av.	Gain	Weighted Gain
GROUP A				
Experimental	25.27	23.5	-1.77	-2.1
Control	27.9	22.5	-5.4	-5.9
GROUP B				
Experimental	26.1	26.	-.1	-.1
Control	26.6	25.5	-1.1	-1.7
GROUP C				
Experimental	85.4	94.16	8.76	7.6
Control	85.92	89.84	3.92	3.1
GROUP D				
Experimental	23.43	24.59	1.16	4.
Control	25.25	26.75	1.50	3.5
GROUP E				
Experimental	50.54	56.	5.46	5.6
Control	49.91	49.1	-.81	-1.2
GROUP F				
Experimental	92.1	96.2	4.1	3.7
Control	91.66	94.9	3.24	2.7
Average all Groups				1.6
Experimental				3.1
Control				.1

TABLE IV

SUMMARY OF SUPERVISORS' RATING OF TEACHERS

	EXPERIMENTAL			CONTROL			DIFFERENCE IN GAIN IN FAVOR OF EXPER.
	1st.	2nd.	Gain	1st.	2nd.	Gain	
PERSONAL QUALITIES							
1	3.45	3.67	.22	2.94	3.44	.50	-.28
2	4.02	4.01	-.01	3.48	4.09	.61	-.62
3	3.19	3.37	.18	2.82	3.16	.34	-.16
4	2.93	3.17	.24	2.59	3.00	.41	-.17
5	3.14	3.35	.21	2.79	3.34	.55	-.34
6	3.27	3.51	.24	3.05	3.47	.42	-.18
7	3.10	3.19	.09	2.53	3.08	.55	-.46
8	3.30	3.61	.31	3.14	3.49	.35	-.04
TOTAL	26.40	27.88	1.48	23.34	27.07	3.73	-2.25
TEACHING SKILL							
1	3.77	4.23	.46	3.74	3.87	.13	.33
2	3.57	3.94	.36	3.25	3.69	.44	-.07
3	2.90	3.70	.80	2.85	3.28	.43	.37
4	2.79	3.86	1.07	2.89	3.14	.25	.82
5	2.62	3.73	1.11	2.65	3.18	.53	.58
6	3.15	3.67	.52	2.80	3.17	.37	.15
7	2.91	3.67	.76	2.80	3.08	.28	.48
TOTAL	21.93	26.86	5.09	20.98	23.41	2.66	2.66
PUPIL REACTION							
1	3.08	3.70	.62	2.73	3.07	.34	.28
2	3.04	3.63	.59	2.80	3.29	.49	.10
3	3.75	4.15	.40	3.45	3.69	.24	.16
4	3.37	3.83	.46	2.96	3.65	.69	-.23
TOTAL	13.24	15.31	2.07	11.95	13.70	1.75	.31
COMBINED TOTAL	61.57	70.05	8.64	56.27	64.18	8.14	.72

TABLE V

SUMMARY OF TEACHER RATINGS, SHOWING AVERAGE INITIAL RATING, FINAL RATING AND GAIN FOR EACH GROUP OF TEACHERS

	PERSONAL QUALITIES			TEACHING SKILL			PUPIL REACTION		
	1st.	2nd.	Gain	1st.	2nd.	Gain	1st.	2nd.	Gain
GROUP A									
Experimental	30.8	35.25	4.45	24.43	30.	5.57	15.72	17.07	1.35
Control	26.70	28.50	1.80	24.70	25.75	1.05	14.86	14.	-.86
GROUP B									
Experimental	28.25	28.58	.33	22.75	24.43	1.68	13.62	13.32	-.3
Control	21.27	28.79	7.52	20.17	24.38	4.21	12.47	15.15	2.68
GROUP C									
Experimental	27.72	30.44	2.72	22.08	28.64	5.56	13.55	15.94	2.39
Control	23.30	25.75	2.45	18.16	19.5	1.34	11.98	11.75	-.23
GROUP D									
Experimental	22.55	24.41	1.86	20.47	24.89	4.42	11.83	13.69	1.86
Control	23.23	24.98	1.75	19.02	21.84	2.82	11.17	12.89	1.72
GROUP E									
Experimental	25.24	22.75	-2.49	21.84	24.07	2.23	12.99	15.25	2.26
Control	23.82	28.75	4.93	24.25	25.5	1.25	10.07	14.	3.93
GROUP F									
Experimental	22.86	27.55	4.69	19.42	27.25	7.83	11.58	16.15	4.57
Control	21.90	26.65	4.75	20.60	24.29	3.69	11.55	14.83	3.28
Average all Groups									1.87
Experimental			1.82			4.55			2.02
Control			3.87			2.39			1.75

"teaching skill", and "pupil reaction." The gain in "personal qualities" was larger for the control group, while in "teaching skill" and "pupil reaction" the experimental group made the improvement. The composite gain is slightly larger for the experimental group.

The two ratings indicate a slight loss by the experimental group in self-control. The control group made its greatest gain in this quality. This factor may be reflected in the fourth item under "pupil reaction"--i.e., group cooperation.

The largest gains made by the experimental teachers were in "questioning quality" and "questioning method." There are but two items under "teaching skill" and "pupil reaction" that indicate a gain favorable to the control groups. These are, "organization of material" under "teaching skill", and "group cooperation", under "pupil reaction", (Table IV). The rating on the latter item, that of "pupil cooperation", seems to contradict the rating on the other items under "pupil reaction." This seems especially true of "attention." The writer has no explanation for the fact that pupils who made the greatest gain in "attention" showed the relatively smaller improvement in "pupil

cooperation." Table VI gives the per cent of the class actually giving attention to the work of the class at 2, 5, 10, 20, and 30 minute intervals of the group conference. The gain is significantly favorable to the experimental group. This same tendency is shown in item 3 of pupil reaction, (Table IV).

TABLE VI

PERCENTAGE OF CLASS GIVING ATTENTION TO THE WORK AT HAND

	EXPERIMENTAL			CONTROL		
Rating Intervals	First Rating	Second Rating	Gain	First Rating	Second Rating	Gain
2 min.	87.8	95.7	7.9	85.5	90.2	4.7
5 min.	86.6	93.6	7.0	87.2	89.3	2.1
10 min	80.6	91.3	10.7	82.8	84.1	1.3
20 min.	75.6	90.2	14.6	81.3	83.2	1.9
30 min.	76.4	87.5	11.1	80.3	83.6	3.3
Average	81.4	91.7	10.3	83.4	86.1	2.7

The attention of the pupils in both groups diminishes distinctly as the class progresses. The diminution is practically the same for all ratings. Teacher "G" of the experimental group is the only individual who had better average attention for the last two intervals than for the first two at both the February and May conferences.

The number of questions asked by teachers and the number of pupil responses during the first ten minutes of the class conference are given in Tables VII and VIII. Both groups reduced the number of questions. The number of pupil responses was increased in the experimental classes while it was correspondingly decreased in the control classes. On the final ratings an average of two pupils in the experimental classes responded to each question, while the control classes averaged less than one response per question.

The results indicate that those groups which received special training in the use of the question as a teaching device made the greater improvement. This was evident in teaching skill, pupil interest, pupil responses, and pupil progress as measured by educational tests. The average gains evidenced by these four criteria were as follows:

	Groups receiving training in questioning	Groups not receiving training in questioning
Teaching Skill	23.2%	12.2%
Pupil Attention	13%	3%
Pupil Response	19.6%	-22.7%
Pupil Achievement (paired groups)	3.1 weighted gain	.1 weighted gain
Pupil Achievement (unpaired groups)	2.53 weighted gain	.25 weighted gain

A significant item in the results of the study is the increase in pupil participation by the experimental group.

TABLE VII

QUESTIONS BY TEACHERS DURING FIRST 10 MINUTES OF CLASS PERIOD

	FIRST RATING			SECOND RATING			DIFFERENCE IN MEAN
	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	
EXPERIMENTAL	2-20	9.9	5.46	3-10	6.2	2.27	-3.65
CONTROL	3-17	10.6	4.47	4-14	8.67	3.4	-2.05
DIFF. IN MEAN		.7			2.47		
S. D. DIFF.		2.			1.1		

TABLE VIII

PUPIL ORAL CONTRIBUTIONS DURING FIRST 10 MINUTES OF CLASS PERIOD

	FIRST RATING			SECOND RATING			DIFFERENCE IN MEAN
	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	
EXPERIMENTAL	3-29	11.2	8.2	8-19	13.4	3.8	+2.06
CONTROL	3-27	11.	6.44	2-21	8.5	4.7	-2.5
DIFF. IN MEAN		.2			4.9		
S. D. DIFF.		2.9					

The increase of 19 per cent in pupil response is accompanied by a higher attention rating and a higher rating on pupil initiative and constructive pupil contribution.

The evidence indicates the cause of this change in pupil activity. While both groups of teachers decreased the number of their questions with experience, only the experimental group increased pupil response. The ratings by experts indicate that the improved pupil response of the experimental group was accompanied by a decided improvement in quality and method of questioning, as well as a reduction in the number of questions; at the same time the control-teachers decreased the number of their questions but showed comparatively little improvement in quality, with a resultant loss in pupil participation.

It would seem that quality rather than number of questions is the index of teaching. If one is to persist in using poor questions, then some advantage accrues from using them rapidly, in so far as pupil response and attention are concerned.

Boardman's Procedures test was administered at the beginning and again at the close of the practice-teaching period. The gain in scores was considered an index of teacher improvement.

On the initial test there was a slight difference in mean, .3, in favor of the control group. On the final test there was a difference favorable to the experimental group of 2.1 with a S. D. Diff. of .8. There was a low correlation between procedures test scores and the ratings of the judges as to teaching skill.

It is difficult to say just what part a knowledge of procedures plays in teaching skill. Undoubtedly one must be able to recognize good practice before he can improve his technique accordingly; but one doubtless knows good technique by sight long before he is able to apply it. It would seem that the gain in score indicates that some knowledge of procedures was acquired, and that the correlation between the score and the rating given the teacher indicates the extent to which this knowledge functioned, --i.e., the effectiveness of the learning.

If this is a sound premise then the change in correlation between skill and recognition-of-procedures would be a valid criterion of the effectiveness of teacher training. If the training is effective the coefficient should increase. In this study, the test scores and teaching skill had a correlation of .07 at the beginning of the practice-teaching period and a correlation of .253 at the close. The second correlation for the experimental group was .34. It was slightly minus for the control group. The experimental teachers not only increased their ability to recognize good teaching practices, but were much more capable of putting into practice those procedures they did know, and this latter is the measure of teaching ability.

Lesson Plans

As an index of teacher-improvement,

lesson plans of student-teachers were evaluated at the beginning and close of the teaching period. Two scales of lesson plans were constructed for the purpose,--one of English plans and one of mixed subjects.

In developing this scale, over four hundred plans were collected from teachers in service and from practice and critic teachers in three teacher-training institutions. From this group, five judges selected representative plans, which were then given rank order ratings by groups of specialists. Preliminary to the selection of the plans for either scale, five separate ratings were made by different groups ranging in numbers from 31 to 44 individuals.

The values of the plans in the scales were determined by the method of converting the per cent-of-better ratings into P. E. differences. From all ratings it was possible to secure two series of thirteen plans with differences in value between each two succeeding plans of .638 to 1.434.

The scales cannot be used jointly since they are based on two sets of ratings. Thus, even though the last specimen of the Beta scale has a P. E. value considerably higher than the corresponding specimen of the Alpha scale, a previous rating which included both of these assigned the two approximately equal value. The values are relative within the scale and in no sense absolute. Anyone desiring to use the scales should choose the one that is most similar to the material to be rated. The Alpha Scale is reproduced at the close of this report.

No attempt was made to determine the validity of the scales. There seems, however, to be little question as to the reliability of the ratings on which the scale was developed. By selecting alternate ratings, totaling all scores, and applying Spearman's formula:

$$P = 1 - \frac{6 \Sigma D^2}{N N^2-1}$$

the correlation was found to be P = .995 for the Alpha scale and P = 1. for the Beta scale. If, instead of totaling the scores, the method of ranking by percentage-of-better ratings is used, as was the case in making the scale, the two sets of judges agree perfectly as to the relative positions of the plans in the scale.

While the quality of lesson plans did not improve in proportion to teaching skill or pupil achievement, the data in this case indicate that ability to make plans, rather than planning, may have been the factor evaluated.

At the beginning and again near the close of the practice-teaching course, as a check on supervisory activities, all practice teachers and supervisors were given a questionnaire in which they were asked to evaluate supervisory method as employed in their group. This questionnaire called for responses to twenty-seven items concerning supervisor-teacher conference, number, type, and length of lesson plans, nature of supervisory assistance during teacher-pupil conferences, and the type of supervisory assistance in selection and organization of materials of instruction.

There is no indication in the data that rather frequent interruptions of a class by the supervisor to correct either physical setting or teaching method need affect either pupil or teacher progress adversely.

The results of the study indicate that supervision by detailed direct criticism is as valuable as supervision by less direct criticism, suggestions, and reference to literature.

While the data collected in this study do not constitute conclusive evidence, they do indicate trends; and conclusions are drawn only where several types of evidence point consistently in the same direction.

SCALE OF LESSON PLANS

Alpha Scale

Sample 1 A - Value 0
English 9
Pages 39-41

Sample 2 A - Value 1.434
English 9 A
Study of Verbs

Sample 3 A - Value 2.247
English 9
Stimulating interest
Writing Letters

Sample 4 A - Value 3.247
English 9
Take up parts of speech, sentences, punctuation, correct expression, capitalization and paragraphing.

Sample 5 A - Value 4.347
English 8
Discussion of parts of speech
Importance of classifying parts of speech
Illustrate the different parts of speech

Sample 6 A - Value 5.160
English 12

Aim: To see that the good novel has or contains some fundamental attitude toward life, some deep seated view of the world in which we live, in the light of which the author sees his characters move and have their being.
"Literature at its best, is a criticism of life." Characters must be people about whom we have opinions. "It is the good reader that makes the good book" (Emerson).

Sample 7 A - Value 6.060
English 7 A
Verbs That Express Action

Aim:
To point out as many verbs that do express action as possible, by having students name verbs - all that they can in the time given.

We shall read according to contract pages 213-214 pointing out verbs of similarity, such as walk (trudges), talk (chatter), spring (jump).
Shall expect at end of hour that students have a better idea of action verbs and have not included such as am, was, etc. Also shall expect them to see that verbs give "life" to sentence and keep action moving.

Sample 8 A - Value 6.698
English 9

Aim: To awaken an appreciation of poetry.
Material:
A Song of Enchantment, by Walter de la Mare.
The West Wind, by James Masefield.

Method:
1. Read A Song of Enchantment.
2. Does de la Mare speak of the present or past?
3. Read The West Wind.
4. Written looking at the past, present, or future?
5. Is regret or longing, in this poem?

Sample 9 A - Value 7.904
English 10

Aim:
To awaken an appreciation of thought and a style of the poem, "The Listeners", by Walter de la Mare.

Procedure:
Discuss briefly with the class the characteristics of lyric poetry and ask them to look for these as we study the poem.
Read the poem to the class.
It is said the best way to enjoy "The Listeners" is to read it over many times.
Ask two or three students in the class to read the poem or parts of it to the class.
Then start a general class discussion using the following questions as a means.
What is the time of day? The

place? Who is the Traveler? Why has he come?

What gives the poem a supernatural element?

What do you hear as you listen to the reading of the poem?

Sample 10 A - Value 8.904
English 7 A

Aim: To show the students that verbs sometimes need help in performing their duties.

Method of Procedure

I. The first ten minutes will be taken to let the class read pages 216-217.
II. How many in the class have ever worked? Had a job or may be helped someone?
III. Do we have to have help from others in order to live? Tell the class what ways you do get help from others.
IV. After having read these pages, what part of speech did you find that needs help the same as you or I?
V. What is an auxiliary verb?
VI. Have some one in the class name the most common auxiliary verbs (ask some member of the class to write them on the board).
VII. Which ones of these are plural and which ones are singular?
VIII. Write these sentences on the board --have them write them in their notebooks, using the form they think is correct.
1. There (was, were) two automobiles going down the street.
2. Those men (has, have) their plans made.
3. The driver and another man (was, were) killed instantly.
4. The boy (have, has) his own coat.
IX. Individual conferences.

Sample 11 A - Value 9.804
English 9

Aim:

To apply material learned on news writing to the play, "Julius Caesar", thus fixing the play more permanently in the minds of the pupils and at the same time getting more practice in news writing.

Material:

1. News items, editorials and feature articles from newspapers.
2. The play, "Julius Caesar."

Method:

1. Review our knowledge of news writing.
2. Write in class news articles whose subjects will be taken from incidents in Julius Caesar.

Procedure:

1. Let us suppose we are living back in the time of Julius Caesar and that we are all writing for a newspaper. Think of the exciting things about which you could write. First of all, however, let us review the material we have learned on good news writing.

2. We learned that a good news item should contain the 5 w's--what are they? Who, what, where, when, why. I also told my class what the object of the news item must be, who can tell me? (To inform.)

Where does the climax of the news item come?

3. Now let us turn to the editorial. What is its purpose? (To make us think and to inform.) Remember, you are trying to make people believe as you believe, but do not let the word "I" enter into the article. Are there any further questions on the editorial or the news item?

4. Here is a list of suggested subjects upon which to write. You need not use these, I shall be glad if you find a topic to write upon yourself.

Act I a. Pretend you are one of the tradesmen to whom Marullus and Flavius talked on the morning of Caesar's return. Write an editorial expressing your attitude toward the celebration.

Act II b. Suppose you were present at the scene which Casca describes wherein the crown was offered three times to Caesar. Write a news item.

Act III c. As an eye witness write a news item on the murder of Caesar.

Act IV d. Write an account of the death of Portia.

Act V e. Suppose you were a war correspondent for a large newspaper. Write an account of the battle between the armies of Octavius and Brutus, of Antony and Cassius.

Write an editorial on the appreciation of Brutus. Include his characteristics, his weakness and his strength. Hand your papers to me at the end of the class.

ASSIGNMENT

Be prepared to hand in your work on contract V on Monday. Miss Weightman and I will be finishing our practice teaching and it will be necessary for your work to be handed to us and corrected next week.

Sample 12 A - Value 10.529
English 10

Aim: To find something in the humorous stories and articles on humor which they haven't noticed.

We always laugh at the unusual or the ridiculous as: Zasu Pitts, Charley Chaplin's shoes, Oliver Hardy and Stanley Laurel, Mickey Mouse, Charles Chase, Louise Fazenda, and Marie Dressler.

What funny commonplace things do we say? For example, on a train going through a tunnel, the porter says, "Look out!" When someone falls down, "Come here, and I'll pick you up."

What things do we laugh at in literature? Do you know any witty people?

Who started American humor and what did he write? Who takes his place today? Which one do you like best?

Why do we like Mark Twain? Does he write of the commonplace, put himself into his books or just use his imagination?

II. What is funny about these expressions that Mark Twain uses in "An Unexpected Acquaintance?" Describe the situation.
1. American.
2. Tail.
3. Names of the children.
4. Age of the girl.
5. I offered my hand, and the party gave me a shake.
6. Seven and a half.
7. Establishment.
8. Sick relative in Siberia.
9. Asylum.

Have you ever wished you had or heard of someone having an experience like this?

III. Long Pants--Irvin S. Cobb.

Did you experience this sensation when you wore your first long pants? Describe his "model's" suit and why it attracted him so much? Did you ever have a model? (I did.) Describe his own new suit. Describe your first suit of long trousers.

Tell about "Speck" and why this name suited him.

Tell about the boy's embarrassment with the girls.

Tell about your own feeling in regard to this.

What did he use at the circus to frighten the people so much? Tell about this incident.

Have you ever wanted to do a prank like this and have you done it?

Tell about the effects of this group.

IV. Correct Behavior on a Picnic--Donald Ogden Stewart.

Why is this called parody and what is a parody? At what is he making fun? Your opinion of etiquette. Your experiences on picnics. Details making this sketch humorous.
1. Family picnic (guest).
2. Describe the start.
3. Describe the guest putting the

top up and down.
4. His conversation.
5. Speed.
6. Entertainment by the guest (suggestion by author).
7. Looking for a place to eat.
8. Their return home.

Don't you think that you have had experiences such as these that are just as funny?

Assignment: Read another humorous story and report on it tomorrow.

Sample 13 A - Value 11.343
English 9

Aim: By impromptu dramatization, to bring "Julius Caesar" before the students as a play which can be acted. Students are prone to believe that this is a story which can only be read.

Materials:

The play, "Julius Caesar", by William Shakespeare.
"A Friend of Caesar", by William Stearns Davis.
"A Day in Old Rome", by William Stearns Davis.

Method:

1. Permit the students to select two out of the following five scenes for dramatic presentation:
   (a) The Ghost Scene.
   (b) The Planning of the Conspiracy.
   (c) Cassius reveals the plot to Brutus.
   (d) Loss of the Battle.
   (e) Death of Brutus.

2. The two scenes having been selected, the students must now decide just what kind of men each of these actors may be.

3. Finally, the class may select the students from their own number who, they think, will suit the parts.

Procedure:

1. Friday, Miss Soupart placed the names of five scenes from Julius Caesar upon the board and requested that each of you read three scenes carefully and select the two which you would like to present. Let us take a vote upon the two scenes which you prefer to present.

2. When you think of George Washington, of what sort of a man do you immediately think? Where did you get your ideas of this man? Most of them were obtained from books you have read about this man. You are able to visualize his character as plainly as though you knew the man himself. So it is with Brutus, Mark Antony, Caesar and all the rest. They too are separate and distinct characters and their characteristics are revealed in this play. Let us see what they are like.

3. First we must make a list of the characters in the scenes which we have selected. One of you pupils may come to the board and as the rest name the characters, he may write them on the board. Let us see how many we can name without our books.

The Conspiracy is planned.
- Casca
- Cicero
- Cassius
- Cinna

Cassius reveals the Plot to Brutus.
- Brutus
- Lucius (servant)
- Cassius
- Casca
- Decius
- Cinna
- Trebonius
- Metellus Cimber
- Portia (wife of Brutus)

Ghost Scene
- Lucius (servant)
- Varro and Claudius (servant)
- Ghost of Caesar
- Messala

Loss of Battle, Death of Brutus
- Clitus
- Cassius, Titinius, Pindarus
- Brutus
- Dardanius
- Volumnius

4. Let us consider Brutus first of all. What line in the play characterizes him best? ("The noblest Roman of them all.") In the ghost scene, Brutus's character is greatly revealed. How? What three things does he ask for? (A book, the gentle boy and music.) Do you think he is a dreamer?

What two things does he place before friendship? (patriotism and pride of race.) Cassius--does he become instigator of the conspiracy because he believes it best for Rome? Why does he? (Envy.) What would you say of his power to suit his argument to the man before him? Where does he make his great mistake? (Surrenders the leadership to Brutus who is an idealist, a dreamer.)

Portia--Why do you think she was introduced? (Reveals more intimate side of husband.) Why do you like her? Do you like her better or less than Calpurnia? Why? Why do you know she was brave? Did Brutus love her? What makes you think thus? Casca, Cinna, Cicero, Metellus, Decius, the other conspirators--What sort of men are these? Why have they joined the conspiracy? Under whose power do we find them?

5. Now, with a complete knowledge of the type of characters needed, let us select the class members whom you think will best play the parts.

6. Let the pupils who have been selected try out reading the parts assigned. The rest of you may be critics. If they do not do so well, perhaps you can improve upon the presentation.

Assignment: Before we can properly reproduce these scenes, we must learn something of the background of the people of the times. We want to know how they lived, what their homes were like and how they dressed.

Let us have some special reports for next time.

In a "Day in Old Rome" read:
Chapt. V, Costume
Chapt. XVI, The Army
Chapt. XVII, The Senate
Chapt. XIX, The Public Games
Chapt. XXII, A Roman Villa

There are also other interesting chapters on Roman food and drink and other subjects. Read them. You will enjoy them.

I have also, here, "A Friend of Caesar", by William Stearns Davis and this too has tales of Caesar. Select something in this book, or any good outside reference book and be able to tell us something either about Roman customs or about Julius Caesar himself.

BIBLIOGRAPHY

[1]Knight, Frederick B. Qualities Related to Success in Teaching. Contributions to Education, No. 120. New York: Bureau of Publications, Teachers College, Columbia University, 1922.

Somers, Grover T. Pedagogical Prognosis. Contributions to Education, No. 140. New York: Bureau of Publications, Teachers College, Columbia University, 1923.

Gambrill, Bessie L., and others. "The Administrative Uses of Intelligence Tests," Twenty-First Yearbook of the National Society for the Study of Education, Part II, pp. 223-243. Bloomington, Illinois: Public School Publishing Company, 1928.

Whitney, F. L. The Prediction of Teaching Success. Journal of Educational Research Monographs, No. 6. Bloomington, Illinois: Public School Publishing Company, 1924.

Tiegs, E. W. An Evaluation of Some Techniques of Teacher-Selection. Bloomington, Illinois: Public School Publishing Company, 1928.

Boardman, Charles W. Professional Tests As Measures of Teaching Efficiency in High School. Contributions to Education, No. 327. New York: Bureau of Publications, Teachers College, Columbia University, 1928.

Morris, Elizabeth Hunt. Personal Traits and Success in Teaching. Contributions to Education, No. 342. New York: Bureau of Publications, Teachers College, Columbia University, 1929.

Ullman, Roy R. The Prognostic Value of Certain Factors Related to Teaching Success. Ann Arbor, Michigan: University of Michigan, 1929. Published by the A. L. Garber Company, Ashland, Ohio.

Breckinridge, Elizabeth. "A Study of the Relation of Preparatory School Records and Intelligence Test Scores to Teaching Success," Educational Administration and Supervision, XVII (December, 1931), 649-60.

Ruediger, W. C., and Strayer, G. D. "The Qualities of Merit in Teachers," Journal of Educational Psychology, I (1910), 272-78.

Bossing, Nelson L. "Teacher Aptitude Tests and Teacher Selection," Research in Higher Education, Bulletin No. 12, pp. 117-135. Office of Education, Washington, D. C., 1931.

Broom, M. E. "A Note on Predicting Success," Educational Administration and Supervision, XVIII (January, 1932), 64-67.

Upshall, C. C., and Masters, Harry V. "An Evaluation of Certain Tests and Information for Predicting Success in Normal Schools," Research in Higher Education, Bulletin No. 12, pp. 21-31. Office of Education, Washington, D.C., 1931.

Meriam, L. L. Normal School Education and Efficiency in Teaching. Contributions to Education, No. 1, pp. 71-80. New York: Teachers College, Columbia University, 1906.

[2]Kriner, Harry Luther. Pre-training Factors Predictive of Teacher Success. Studies in Education, No. 1. State College, Pennsylvania: Pennsylvania State College, 1931.

[3]Barr, A. S., An Introduction to the Scientific Study of Classroom Supervision, p. 276f. New York: D. Appleton and Company, 1931.

[4]Kyte, G. C. Reorganization and Administration of an Elementary School to Meet the Needs of a Community, pp. 58-59. Berkeley, California: University of California, 1922. Review given in Fourth Yearbook, Department of Supervisors and Directors of Instruction, National Education Association, pp. 58-59, 1931.

[5]Pittman, M. S. The Value of School Supervision. Baltimore, Maryland: Warwick and York, Inc., 1921.

[6]Collings, Ellsworth. An Experiment with a Project Curriculum. New York: The Macmillan Company, 1923.

[7]Hoppes, W. C. The Value of Supervision in the Rural Schools of Oakland County. Bulletin No. 7. Lansing, Michigan: Michigan Education Association, 1926.

[8]Sherwood, H. W. Value of Rural School Supervision. Educational Bulletin No. 84, Indiana State Department of Public Instruction. Indianapolis, Indiana, 1926.

[9]Brueckner, Leo J. The Technics and Evaluation of a Supervisory Program in Work Reading in the Minneapolis Public Schools. Educational Bulletin, No. 12. Minneapolis, Minnesota: Minneapolis Public Schools, 1927.

[10]Melby, Ernest O. Organization and Administration of Supervision. Northwestern University Contributions to Education, School of Education Series, No. 1. Bloomington, Illinois: Public School Publishing Company, 1929.

[11]Valentine, P. F. "A Job Analysis of Elementary Supervision," Journal of Educational Method, V (March, 1926), 279-82.

[12]Kyte, G. C. How to Supervise, p. 400f. Boston: Houghton Mifflin Company, 1930.

[13]Leman, Grant William. A Study of Factors in the Development of Professional Adjustment Service. New York: New York University, 1932.

[14]Pickett, Lalla H. An Analysis of In-service Training Programs of Twenty-five Selected Normal Schools and Teachers Colleges. New York: New York University, 1932.

[15]Fuda, Anna M. Teacher Judgment of In-service Education. In-service Education Section of the National Survey of the Education of Teachers. New York University, 1932.

[16]Fuller, Henry Harrison. The Relative Effectiveness of the Group Supervisory Conference and the Individual Supervisory

Conference. Ann Arbor, Michigan: University of Michigan, 1929.

[17]Marshall, Edna M. Evaluation of Types of Student Teaching. Contributions to Education, No. 488. New York: Bureau of Publication, Teachers College, Columbia University, 1932.

[18]Jarman, Arthur Merritt. The Administration of Laboratory Schools. Ann Arbor, Michigan: George Wahr, 1932.

[19]Flowers, John Garland. Content of Student Teaching Courses Designed for the Training of Secondary Teachers in State Teachers Colleges. New York: Bureau of Publications, State Teachers College, Columbia University, 1932.

[20]Voelker, Otto H. Provision for Laboratory School Experience in New York State. New York University, 1932.

[21]Learned, William S., Bagley, W. C., and others. The Professional Preparation of Teachers for American Public Schools. Bulletin No. 14. New York: The Carnegie Foundation for the Advancement of Teaching, 1920.

[22]Jarman, Arthur Merritt. op. cit., p. 120.

[23]Fourth Yearbook, American Association of Teachers Colleges, p. 111. Washington, D. C.: National Education Association, 1925.

[24]Armentrout, W. D. The Conduct of Student Teaching in State Teachers Colleges. Colorado State Teachers College. Education Series, No. 2. Greeley, Colorado, 1927.

[25]Hudelson, Earl. Educational Supervision. First Yearbook, National Conference on Educational Method. New York: Bureau of Publications, Teachers College, Columbia University, 1928.

[26]Rankin, Paul T., and others. Scientific Method in Supervisory Programs. Seventh Yearbook, Department of Supervisors and Directors of Instruction, National Education Association, p. 22. New York: Bureau of Publication, Teachers College, Columbia University, 1924.